EVALUATION OF DIAGNOSTIC TESTS FOR DETECTION OF MYCOBACTERIUM AVIUM SUBSP. PARATUBERCULOSIS (MAP) AT THE HERD-LEVEL AND COW-LEVEL

BY

CARRIE J. LAVERS

A Thesis
Submitted to the Graduate Faculty
In Partial Fulfillment of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

Department of Health Management
Faculty of Veterinary Medicine
University of Prince Edward Island

CONDITIONS FOR THE USE OF THE THESIS

PERMISSION TO USE GRADUATE THESIS

Title of Thesis:

EVALUATION OF DIAGNOSTIC TESTS FOR DETECTION OF MYCOBACTERIUM AVIUM SUBSP. PARATUBERCULOSIS (MAP) AT THE HERD-LEVEL AND COW-LEVEL

Name of Author: Carrie J. Lavers

Department: Health Management

Degree: Doctor of Philosophy Year: 2013

Name of Supervisor(s): Dr. Greg Keefe and Dr. Shawn McKenna

In presenting this thesis in partial fulfilment of the requirements for a graduate degree from the University of Prince Edward Island, the author has agreed that the Robertson Library, University of Prince Edward Island, may make this thesis freely available for inspection and gives permission to add an electronic version of the thesis to the Digital Repository at the University of Prince Edward Island. Moreover the author further agrees that permission for extensive copying of this thesis for scholarly purposes may be granted by the professor or professors who supervised the author's thesis work, or, in their absence, by the Chair of the Department or the Dean of the Faculty in which the author's thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without the author's written permission. It is also understood that due recognition shall be given to the author and to the University of Prince Edward Island in any scholarly use which may be made of any material in the author's thesis.

Signature: _____

Address: 550 University Avenue, Charlottetown, PEI, C1A 4P3

Date: _____

University of Prince Edward Island

Faculty of Veterinary Medicine

Charlottetown

CERTIFICATION OF THESIS WORK

We, the undersigned, certify that **Carrie J. Lavers, DVM**, candidate for the degree of Doctor of Philosophy has presented her thesis with the following title:

EVALUATION OF DIAGNOSTIC TESTS FOR DETECTION OF MYCOBACTERIUM AVIUM SUBSP. PARATUBERCULOSIS (MAP) AT THE HERD-LEVEL AND COW-LEVEL,

that the thesis is acceptable in form and content, and that a satisfactory knowledge of the field covered by the thesis was demonstrated by the candidate through an oral examination held on December 17, 2013.

Examiners:

_____ Dr. Jeff Davidson

_____ Dr. Spencer Greenwood

_____ Dr. Liz Spangler

_____ Dr. John VanLeeuwen

_____ Dr. Scott Wells

Date: _____

ABSTRACT

The purpose of this research was to evaluate herd and cow-level Mycobacterium avium subspecies paratuberculosis (MAP) diagnostic tests within non-infected and low to moderate within-herd prevalence MAP-infected herds. Specific objectives were to estimate herd-level sensitivity and specificity of environmental culture and milk ELISA, determine predictors influencing herd test characteristics, evaluate cow-level test characteristics and between-test agreement of commercial milk ELISAs, and quantify additional diagnostic information gained from repeating a milk ELISA test.

A longitudinal study was conducted between May 2009 and February 2011 on 34 dairy herds from Prince Edward Island, Nova Scotia, and New Brunswick. Environmental manure samples were collected every 3 mo, and cow manure and milk samples were collected from all lactating cows in the herd every 6 mo. Pooled fecal culture was used to establish MAP herd status. Individual fecal culture from cows within positive fecal culture pools was used to determine apparent within-herd MAP prevalence. Cow and environmental manure samples were cultured in a broth system, with confirmatory PCR targeting the hspX gene. Milk samples were analyzed using 3 commercial milk ELISAs. For herd-level statistical analyses, pseudogold standard methods were applied (herd reference: repeated pooled fecal culture). For cow-level test characteristics, pseudogold standard methods (reference: fecal culture) and latent class analyses were performed. Generalized estimating equation models accounted for repeated measures.

Herd sensitivity of environmental culture was 71%, and specificity was 99%. Herd sensitivity of 3 milk ELISAs ranged from 56 – 63%, and herd specificity from 80 –

96%. Herd sensitivity of both environmental culture and milk ELISA improved as within-herd MAP prevalence increased. However, environmental culture sensitivity was higher at lower within-herd prevalence than herd milk ELISA tests, indicating environmental culture is a more appropriate test to detect low-prevalence herds. The proportion of positive samples within an environmental culture set was positively associated with within-herd MAP prevalence, establishing that environmental culture provides an estimation of within-herd prevalence for further risk management decisions.

Evaluation of cow-level milk ELISA results revealed between-test agreement was low for positive ELISA tests. Cow-level milk ELISA sensitivities ranged from 28 – 35%, and specificities from 99 – 100%, and were not different for pseudogold standard methods or latent class models. Likelihood ratios indicated that cows with ELISA results in the uppermost quantitative category had a greater likelihood of MAP fecal shedding, necessitating specific management strategies for these cows. When a milk ELISA test was repeated 12 mo later, sensitivity increased from 26% for the initial test only, to 45% for the combined tests, while specificity decreased minimally. A repeated milk ELISA test substantially improves the probability to detect a MAP-infected cow, and should be considered for a herd control program.

Development of effective herd and cow-level paratuberculosis testing programs is essential for the implementation of a successful MAP control program. Our herd-level analyses provide an estimation of the impact of within-herd prevalence on herd test characteristics, which will assist in selection of appropriate herd tests. Our cow-level ELISA analyses provide test characteristics of commercially available ELISAs, and extend the use of ELISA results beyond a single dichotomous outcome. Consideration of

our quantitative ELISA results and interpretation of repeat ELISA tests will further improve management recommendations for the development of successful MAP control programs.

ACKNOWLEDGEMENTS

Many people have helped me over the last four years, and without them this thesis would not be possible, nor would the journey have been nearly as much fun.

I am indebted to my co-supervisors, Drs. Greg Keefe and Shawn McKenna, for their support. Greg has always managed to see the positive side of, and find a solution for, the challenges we encountered. Despite a very full schedule, he has always made time to help and reassure me when it was needed, and I appreciate that no call for help went unanswered. Shawn has also been a supportive and practical voice, providing advice, encouragement, and humor throughout this project. Shawn and Greg have both been understanding of the demands on my time that have come from outside of work, and for that I am extremely grateful.

I was fortunate to have an excellent supervisory committee: Drs. Herman Barkema, Spencer Greenwood, J McClure, Henrik Stryhn, and John VanLeeuwen. I am thankful they took time out of their busy schedules to help me through this program. From initial planning to final thesis revisions, their advice was invaluable.

I am so thankful to "Team Johne's" for working hard, and making sample collection enjoyable: Theresa Andrews, Heather Jack, Jane Saunders, Ross Campbell, and Lindsay MacNeil. Theresa deserves extra-special recognition for cheerfully doing a myriad of additional tasks that made the collection of samples, and the reporting of results back to veterinarians, run smoothly.

This project involved an immense amount of laboratory work, and I am extremely grateful for the organized efforts of the talented staff at the Maritime Quality Milk Laboratory: Natasha Robinson, Dr. Marcelo Chaffer, Maria Vasquez, Kate McQuillan, Doris Poole, as well as numerous summer students. In particular, thanks to Natasha for

ensuring the laboratory processes were done properly and efficiently, and for calmly and quickly answering my countless lab-related queries.

I am very grateful to the farmers who generously allowed us to collect samples from their cows and barns. I know that every minute is precious on a farm, and they surely had many other jobs they could have been doing, yet they shared their time and their cows so that we could learn. It is producers like these that will keep the dairy industry growing and progressing.

Drs. Ian Dohoo and Henrik Stryhn played a significant role in the analysis of data for this thesis, and I could not have done this without them. They both made me feel that every question was reasonable, and were able to clearly explain complicated methods. I always left their offices amazed at their brilliance, lifted up because an analysis made more sense, and wishing I could record our conversations for playback later. The graduate students are truly blessed to have two such amazing teachers that are willing to share their talents and knowledge. I would also like to thank the following staff members that helped with portions of data management for the analysis: Jeff Carmody, Elmabrok Masaoud, and Jenny Yu.

Fellow graduate students were always important sources of information and encouragement: Drs. Raph Vanderstichel, Charles Caraguel, Kim MacDonald, Kristen Reyher, Dave McIver, Colleen Walton, and Patti Jones. In particular, I was lucky to have Dr. Maggie Cameron at the desk beside me, to keep me company, and to keep me laughing. Thanks also to Drs. Emilie Laurin and Karen MacDonald-Phillips, fellow Johne's grad students, for sharing data, papers, and an interest in MAP.

And finally, thank you to my husband, Mike Carruthers, for his continuous support over these last four years. I appreciate all the nights that he took over the seemingly endless jobs that are involved in raising a young family, so that I could get just a little more work done.

DEDICATION

This thesis is dedicated to my husband and children, who gave up many date-nights and bedtime snuggles, in order that this thesis could be completed,

And to Mark Robert Adams, whose name would have been on the cover of a doctoral thesis in Philosophy, but for the universe setting a different stage.

“Change is always inherently good, and growth always happens, even if it’s outside of our eyes.”

TABLE OF CONTENTS

List of Tables

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 6

List of Figures

CHAPTER 2

CHAPTER 3

CHAPTER 4

List of Abbreviations

d	day(s)
DHI	dairy herd improvement
EC	environmental culture
ELISA	enzyme-linked immunosorbent assay
FC	fecal culture
GEE	generalized estimating equation
HSe	herd sensitivity
HSp	herd specificity
IgG	immunoglobulin G
IFN-γ	interferon-gamma
JD	Johne's disease
LR	likelihood ratio
MAP	Mycobacterium avium subspecies paratuberculosis
min	minute(s)
mL	millilitre
mo	month(s)
OD	optical density
OR	odds ratio
PCR	polymerase chain reaction
PFC	pooled fecal culture
PFC-IC	pooled fecal culture, followed by individual fecal culture of samples from culture positive pooled fecal cultures
Se	sensitivity
Sp	specificity
S/P	sample to positive ratio
US	United States

WHP	within-herd prevalence
yr	year(s)
µg	microgram

CHAPTER 1. GENERAL INTRODUCTION

Paratuberculosis (Johne's disease) is a chronic, infectious enteritis of dairy cattle caused by Mycobacterium avium subspecies paratuberculosis (**MAP**) (Chiodini et al., 1984). Paratuberculosis has long been recognized by the veterinary community as a serious disease affecting cattle. The first report of this condition dates back to 1826, and through the efforts of illustrious scientists such as Johne, Frothingham, Koch, and Bang, came to be known as Johne's disease in 1905 (Chiodini et al., 1984). Despite a disease history stretching back almost 200 years, MAP continues to challenge researchers in regards to diagnosis, treatment, and control.

Paratuberculosis has been reported worldwide (Collins, 2003; Singh et al., 2013), and the natural hosts are domesticated and wild ruminants, including dairy cattle, beef cattle, sheep, goats, cervids, camelids, and red deer (Kennedy and Benedictus, 2001). There have also been reports of a wide range of non-ruminant wildlife species carrying MAP, including, but not limited to, rabbits (Raizman et al., 2005), rhinoceros (Bryant et al., 2012), equines, primates, rodents, and birds (Münster et al., 2013). Although these non-ruminant species represent risks as potential MAP reservoirs, proof of transmission is still required (Hutchings et al., 2010).

1.1 Mycobacterium avium subspecies paratuberculosis

The MAP organism is a member of the Mycobacterium avium complex, which is divided into 2 species of slow-growing, nontuberculous mycobacteria: M. intracellulare and M. avium (Pfyffer, 2006). The M. avium species is further divided into 4 subspecies, M. avium subsp. paratuberculosis, M. avium subsp. avium, M. avium subsp. silvaticum,

and M. avium subsp. hominissuis (He and De Buck, 2010). The MAP organism is differentiated from other members of the M. avium complex by the presence of unique genes, including IS900 (Green et al., 1989), and hspX (Ellingson et al., 1998).

Mycobacterium avium subsp. paratuberculosis is an aerobic, acid-fast bacillus, and is 0.5 – 1.5 µm long (Clarke, 1997). In culture, MAP is slow-growing, with a doubling time of 22 – 26 hr, compared to 10 – 12 hr for M. avium subsp. avium (Bannantine et al., 2003). The bacterium is fastidious to culture because, unlike most mycobacteria, MAP does not produce iron-chelating mycobactin, and therefore requires an exogenous mycobactin source when cultured (Clarke, 1997). The organism has a hydrophobic, thick, and waxy outer cell well (Pfyffer, 2006), which contributes to the organism's antibiotic resistance, and its ability to survive in a variety of environmental conditions (He and De Buck, 2010).

The MAP organism is an obligate intracellular pathogen, and cannot replicate outside the host (Harris and Barletta, 2001; Bannantine and Stabel, 2002). However, it is able to survive for up to one year in the environment if it is protected from sunlight (Whittington et al., 2004). While mycobacteria are typically sensitive to ultraviolet light (sunlight), they are not easily destroyed by heat, freezing or desiccation, acid or alkaline compounds, or chemical agents (Pfyffer, 2006).

1.2 Transmission

It is a commonly held assumption that cattle are primarily infected with MAP as calves, and become more resistant to infection as they age (Hagan, 1938; Taylor, 1953; Larsen et al., 1975; Chiodini et al., 1984). A meta-analysis of infection studies concluded that there was a significant difference in age susceptibility to infection between adults

and calves less than 6 mo of age and between adults and calves 6 – 12 mo of age (Windsor and Whittington, 2010). Evidence from experimental infection studies suggest that the proportion of successfully infected calves is expected to be 75% in calves < 6 mo, 50% in calves 6 – 12 mo, and 20% in cattle > 12 mo (Windsor and Whittington, 2010). A recent calf infection study confirmed that calves up to 1 yr of age were more susceptible to MAP infection, although calves inoculated at < 6 mo of age typically had more culture-positive tissue locations (Mortier et al., 2013). Although cows can be infected with MAP when exposed to the bacterium as an adult, they are less likely than exposed calves to be seropositive or fecal culture-positive, or to be culled from the herd due to clinical signs of paratuberculosis (Espejo et al., 2013). Data from previous studies are considered insufficient to determine the age at which susceptibility diminishes in adults, compared to calves (Windsor and Whittington, 2010).

The most relevant natural route of infection with MAP is considered to be ingestion of the bacterium in fecal-contaminated feed, milk, colostrum or water (Stabel, 1998; Tiwari et al., 2006; Begg and Whittington, 2008). Milk can also be contaminated with MAP due to heterogeneous spread of the bacterium into the milk of infected cows, with a higher prevalence of milk infection occurring in heavy fecal shedding cows, compared to cows shedding low levels of MAP (Streeter et al., 1995; Sweeney et al., 1992a). Contaminated environments may also be sources from which the bacterium might be ingested. The MAP bacterium percolates slowly through soil, and remains on grass and in the upper soil levels, representing a risk of ingestion for grazing cattle (Salgado et al., 2011). The organism can survive for 1 year in the water and sediment of shaded water troughs (Whittington et al., 2005). The bacterium has also been detected in

biofilms, including those present in cattle watering troughs (Cook et al., 2010). In addition, viable MAP has been detected in settled dust in dairy barns, and it has been suggested that bioaerosols may pose a MAP infection risk (Eisenberg et al., 2010; Eisenberg et al., 2012). Furthermore, a calf infection study indicated that inhalation of aerosolized MAP can result in infection (Eisenberg et al., 2011).

Another documented route of infection is vertical transmission from an infected dam to a fetus in utero (Seitz et al., 1989; Sweeney et al., 1992b). A review and meta-analysis identified that vertical transmission is a relevant concern that should be addressed in herd control programs (i.e. calves from cows identified infected by histology or microbiology should be removed from the herd) (Whittington and Windsor, 2009). However, the review also recognized the need for further research in order to better understand and intervene in the route of in utero transmission (Whittington and Windsor, 2009).

Infection studies have demonstrated that a higher infection dose produces more severe tissue lesions than a lower dose (Begg and Whittington, 2008; Mortier et al., 2013), as well as a shorter lag phase between infection and an observed immunological response (O'Brien et al., 2006). The greater the dosing frequency, the less likely a calf will mount an adequate immune response to clear MAP infection in the interval between doses (Begg and Whittington, 2008). The exact concentration of MAP, and number and frequency of oral doses, required to result in MAP infection is unknown, and these parameters have varied greatly in infection studies (Begg and Whittington, 2008). However, results of these infection studies suggest that the higher the MAP burden and

the more frequently the calf is exposed to it, the more likely that MAP infection will occur.

1.3 Host Response

After ingestion, MAP organisms enter ileal tissues through M cells, specialized nonvillous epithelial cells in the Peyer's patches of the ileum (Momotani et al., 1988). After crossing the intestinal epithelial layer, the organism is phagocytosed by sub-epithelial macrophages (Singh et al., 2013). The bacterium resists degradation by residing and replicating within immature phagosomes, inhibiting phagosomal maturation processes that normally result in destruction of intracellular bacteria (Hostetter et al., 2003). In addition, MAP decreases macrophage apoptosis, delaying the immune system's exposure to, and recognition of, the bacterium (Kabara and Coussens, 2012). Intracellular survival of MAP within these macrophages is key to its pathogenicity (Hostetter et al., 2003). When the macrophage's phagosomes are able to process the bacterium and present it to T lymphocytes, the immune process begins (Stabel, 2000).

Initially, a cell-mediated immune response predominates, with production of T-helper lymphocytes, which produce the cytokine interferon-gamma (**IFN-γ**) (Stabel, 2000). This cell-mediated response is a critical defense against MAP infection, and is usually mounted prior to the development of antibody production (Chiodini, 1996). In some cows, the cell-mediated immune response may be adequate to control or eradicate MAP infection (Mikkelsen et al., 2009). In cows in which cellular immunity is unable to control infection, the cell-mediated response diminishes and the humoral response begins to predominate (Stabel, 2000), with the production of non-protective antibodies and cytokines (interleukin-4 and interleukin-10) (Sweeney, 2011). In cattle, the predominant

antibody produced in response to MAP infection is immunoglobulin G (**IgG**) (Abbas and Riemann, 1988).

1.4 Pathogenesis

1.4.1. Silent Infection

In the early stages of infection, host defenses are able to contain the MAP infection, and only slow proliferation and spread of MAP within the ileal submucosa and mesenteric lymph nodes occurs (Sweeney, 2011). This stage of infection, which generally lasts for 2 or more years, is often termed silent infection because the animal shows no clinical or subclinical signs of infection, and fecal shedding of the bacterium and an antibody response are not yet detectable (Whitlock and Buergelt, 1996). Although the MAP infection is being controlled and contained at this stage, MAP antigens within the intestinal submucosa and mesenteric lymph nodes elicit an anti-inflammatory response, as the host attempts to contain the infection. As additional macrophages and lymphocytes are attracted to the area of infection, granuloma formation occurs (Sweeney, 2011).

1.4.2. Subclinical Infection

Although the granulomatous inflammation contains the MAP infection for a period of time, eventually the infection overtakes the immune response (Sweeney, 2011). This appears to happen at the same time as the cell-mediated immunity wanes and the humoral immune response takes over (Stabel, 2000), and at this point, infection progresses more rapidly (Sweeney, 2011). As the immune response starts to lose control of the infection, MAP is shed in increasing quantities in the feces, and the MAP organism spreads to other tissues in the body. This stage is termed subclinical infection, because

although the cow typically does not show clinical signs, subclinical effects begin to manifest (Whitlock and Buergelt, 1996). Potential subclinical manifestations of paratuberculosis include decreased milk production (Nielsen et al., 2009) and infertility (Stabel, 1998).

Detection of MAP infection is challenging at this stage. Fecal shedding of the bacterium can be intermittent (Nielsen, 2008), and antibody development may not yet be detectable (Sweeney, 2011). Literature results suggest that a detectable antibody response may occur either after detection of bacterial shedding (Lepper et al., 1989; Sweeney et al., 2006), or before detection of bacterial shedding (Nielsen, 2008). This variation between time to occurrence of bacterial shedding and occurrence of antibody response may be related to the disparity in infectious dose the calf is exposed to (Nielsen and Toft, 2008). There is a correlation between bacterial shedding and antibody (IgG) response, and cows shedding MAP generally have higher concentrations of IgG antibodies (Koets et al., 2001). The variation in time to occurrence of either response, as well as the intermittent nature of these host responses, translates to poor diagnostic test sensitivity during this stage. The practical implication is that during this stage, infected cows more frequently test negative than at later stages. However, these false-negative cows may shed low numbers of MAP in their feces, contaminating the herd environment and posing an infection threat to susceptible animals (Tiwari et al, 2006). The duration of subclinical fecal shedding is variable, and some cows show clinical signs within 6 mo of detectable fecal shedding, while others may shed MAP in their feces for years without demonstrating clinical signs of infection (Sweeney, 2011).

1.4.3. Clinical and Advanced Clinical Infection

As granulomatous lesions in the jejunum, ileum, cecum, and to a lesser extent, colon, as well as mesenteric lymph nodes, become more severe, the lining of the small intestine (particularly the ileum) becomes thickened due to the cellular infiltration (Stabel, 1998). This inflammation results in malabsorption, diarrhea, and a protein losing enteropathy (Stabel, 1998). These pathological changes lead to the clinical and advanced clinical signs of the disease, which are progressive weight loss, profound watery diarrhea, and diffuse edema (Whitlock and Buergelt, 1996). Age at onset of clinical signs is variable, most frequently considered to range from 2 to 5 yr (Larsen et al., 1975), although onset of clinical signs has been noted in cows up to 10 yr of age (Whitlock and Buergelt, 1996). During the clinical and advanced clinical stages, most cows demonstrate both an antibody response and bacterial shedding (Tiwari et al., 2006). Only 10 to 15% of infected animals reach this clinical stage of paratuberculosis, because the majority are culled due to productivity losses in the earlier subclinical stages of disease (Abbas et al., 1983).

1.5 Diagnostic Testing and Target Condition

Available ante-mortem paratuberculosis tests are based on either detection of the organism (fecal culture and fecal polymerase chain reaction (**PCR**)), or detection of an immune response (IFN-γ or IgG). Fecal culture is considered the most sensitive and specific ante-mortem paratuberculosis diagnostic test (Whitlock et al., 2000). However, compared to other available cow-level tests, it is the most costly (e.g. $35 to $60), and takes the longest to complete (12 to 16 wk) (Tiwari et al., 2006). Fecal PCR is a much faster laboratory procedure than culture, with a 4 d turnaround (Tiwari et al., 2006). There have been considerable improvements to the procedure in the last decade, and the

sensitivity may approach that of fecal culture, although fecal culture remains the definitive reference standard in research applications (Bölske and Herthnek, 2010).

Immune-based diagnosis of paratuberculosis includes detection of a cell-mediated immune response, or more specifically, IFN-γ. Because some animals with a positive IFN-γ response may be able to control and eradicate MAP infection, this test is used only to determine if an animal has been exposed to MAP (Mikkelsen et al., 2009). Finally, antibody (IgG) levels can be measured in serum and milk using enzyme-linked immunosorbent assays (**ELISA**), which measure an optical density level that correlates to the antibody response. These tests have a quick turnaround time (1 d) and are cost-effective (e.g. $10) (Tiwari et al., 2006). It should be noted that costs of diagnostic tests vary from laboratory to laboratory, and costs provided here are intended to provide an example of the difference in costs between fecal culture and milk ELISA. They are not inclusive of all possible testing costs.

When diagnostic tests are evaluated, it is important to establish the target condition (Nielsen and Toft, 2008), which reflects the underlying MAP status being detected (affected, infectious, or infected cows or herds) (Gardner et al., 2011). Affected cows demonstrate at least 1 clinical sign of paratuberculosis (Nielsen and Toft, 2008). Infectious animals shed MAP, and are a risk for infection to susceptible herdmates (Nielsen and Toft, 2008). This target condition also includes cows that are affected because affected cows are also shedding MAP (Nielsen and Toft, 2008).

Infected animals carry MAP intracellularly (Nielsen and Toft, 2008). In an animal infected with MAP, MAP infection has persisted to the point that the cow may

demonstrate an immunological response at any point during her life, and it is assumed that once infection is established, it persists for life (Nielsen and Toft, 2008). The infected target condition also includes affected and infectious cows because affected and infectious cows are clearly infected.

At the herd level, Gardner et al. (2011) recognized an infected herd as a target condition, which is a herd that has at least 1 infected cow. The Gardner et al. (2011) consensus document does not include infectious or affected target conditions for herds. Presumably, an infectious herd has at least 1 infectious cow, and an affected herd has at least 1 affected cow.

Relative to the four stages of disease discussed earlier (silent, subclinical, clinical, and advanced clinical stages of disease), the affected target condition includes cows in the clinical or advanced clinical stages of disease. The infectious target condition encompasses cows that are in the subclinical, clinical, and advanced clinical stages of disease. Finally, the infected target condition could exist within any of the 4 disease stages that describe paratuberculosis.

1.6 Economic Impact

Productivity losses are a key factor in the economic cost of paratuberculosis. Some of the more substantial impacts on productivity include decreased milk production (Benedictus et al., 1987; Wilson et al., 1993; Nordlund et al., 1996; Lombard et al., 2005; Tiwari et al., 2007; Nielsen et al., 2009), increased risk of culling (Wilson et al., 1995; Lombard et al., 2005; Tiwari et al., 2008), decreased slaughter value (Benedictus et al., 1987; Chi et al., 2002), and decreased slaughter weight (Johnson-Ifearulundu et al., 1999). These production losses have an economic cost of $200-250 million annually to

the United States dairy industry (Ott et al., 1999). Direct economic impact to the Canadian Maritime provinces has been conservatively estimated at $0.84 million annually (Chi et al., 2002), with a projected cost of $15 million annually to the whole of the Canadian dairy industry (McKenna et al., 2006b). To the author's knowledge, there are not more recently published estimates of the economic impact of paratuberculosis in Canada or the United States. Potentially, the economic burden is now greater, due in part to the realities of inflation and rising costs of dairy production, and also because herd-level prevalence of paratuberculosis has increased over the last decade (Lombard et al., 2008).

In addition to the substantial production limiting impacts of the disease, a further motivation to diagnose infection and control the spread of MAP is the concern of a possible association between the bacterium and Crohn's disease in humans (Sanderson et al., 1992; Bull et al., 2003; Naser et al., 2004). Systematic reviews and meta-analyses have concluded that, while there does appear to be an association between MAP and Crohn's disease, further work is required to fully understand the pathogenesis of any such association (Feller et al., 2007; Abubakar et al., 2008; Barkema et al., 2011). Although the zoonotic potential of MAP remains uncertain, it is paramount that the dairy industry makes a full effort to minimize human exposure to the organism.

1.7 Prevalence

The majority of animal-level, within-herd and herd apparent prevalence estimates available from large scale, randomized studies are based on serum ELISA testing. Because ELISAs have a low sensitivity (McKenna et al., 2005), these estimates are likely underestimates of the true prevalence. In a seroprevalence study of 90 randomly selected

herds in the Canadian Maritime provinces, 2.6% of cows were seropositive (animal-level prevalence), and 17% of herds were MAP-positive (herd prevalence). Within MAP seropositive herds, the average within-herd test prevalence was 8.5% (VanLeeuwen et al., 2001). In a seroprevalence study of 1,004 herds in the United States, 3.4% of cows (animal-level prevalence) and 22% of herds (herd prevalence) were MAP-seropositive (Wells and Wagner, 2000). More recently, the Ontario Johne's Education and Management Assistance Program has, to date, performed whole herd milk ELISA testing in 2,215 herds in Ontario, Canada. Twenty-six percent of these herds had at least 1 ELISA-positive cow (herd prevalence). Of the 146,704 cows tested, 1% were milk ELISA-positive (animal-level prevalence) (Ontario Johne's Education and Management Assistance Program, 2013).

When looking at organism detection methods, composite environmental fecal samples have been used to estimate that 68% of American dairy herds were MAP-positive (USDA, 2008). These US data have recently been re-analyzed using Bayesian methods to correct for test characteristics, and a herd prevalence of 91% was projected (Lombard et al., 2012). The authors noted this estimate was not comparable to previous studies due to different testing methods and herd classification criteria. Comparison between prevalence studies is often limited because of such differences. Despite variation in estimates, paratuberculosis is generally considered an infection with a low to moderate within-herd prevalence (Lombard, 2011).

1.8 Control Programs

There is not an approved vaccine for paratuberculosis in Canada, and only 1 licensed in the United States (Patton, 2011). While the vaccine decreases clinical disease

and fecal shedding, it does not prevent infection, and has not been widely used (Patton, 2011). Use of the vaccine is further limited because a permanent granulomatous lesion often develops at the vaccination site (Larsen et al., 1978). There is also no treatment that definitively cures paratuberculosis, and no treatment approved for use in food-producing cattle (Fecteau and Whitlock, 2011). There are therapeutic agents that will reduce or alleviate clinical signs, but these must be given for the life of the animal, and are typically only used on animals with substantial economic or genetic value (Fecteau and Whitlock, 2011).

Given that there is no effective treatment or vaccination available to cure or prevent paratuberculosis infection, management of the disease relies on successful MAP control programs (Garry, 2011). The United States' National Voluntary Bovine Johne's Disease Control Program suggests 3 sequential steps for a MAP control program: producer education, followed by development of a risk assessment and management plan, and finally establishment of a cow-level testing program (Whitlock, 2010).

Before developing a MAP control program for a herd, it is necessary to establish whether or not a herd is infected, and obtain an estimate of the MAP within-herd prevalence (Garry, 2011). Using this herd-level information, development of a MAP control program can proceed. The control program is individualized for a herd, by taking into consideration factors such as MAP herd-status and within-herd prevalence, as well as herd risk factors identified during development of a herd risk assessment and management plan. The last step in a MAP control program, for MAP-infected herds, is to develop an ongoing cow-level testing program. The purpose of cow-level testing is to identify cows that are most infectious (Garry, 2011), and when test results are used to

initiate management changes for MAP-infectious cows, the return on investment is usually justified (Dorshorst and Lombard, 2006).

1.9 Limitations in Paratuberculosis Research

Despite the expectation that the majority of MAP-infected herds experience low to moderate within-herd prevalence (Lombard, 2011), many paratuberculosis studies evaluating herd and cow-level test characteristics have selected known MAP-positive herds, often having high within-herd prevalence, as the study population. While this strategy maximizes the use of valuable research dollars, external validity can be compromised. At the herd-level, test characteristics are influenced by the true within-herd prevalence (Christensen and Gardner, 2000). At the cow-level, the impact of within-herd MAP prevalence on cow-level test characteristics is not as well documented. However, statistical modeling suggests that disease prevalence can cause variation in sensitivity and specificity estimates made from different populations (Brenner and Gefeller, 1997). Greiner and Gardner (2000) also note increasing evidence that, for many tests, the sensitivity and specificity vary with the characteristics of the population to which they are applied. It is therefore challenging to extend herd and cow-level test characteristics, developed in high-prevalence study herds, to target herds with low to moderate within-herd MAP prevalence.

Another limitation in previous paratuberculosis research is the infrequent use of a longitudinal study design to establish a case definition for the test under evaluation. Repeated sampling maximizes identification of MAP-infected low prevalence herds (Wells et al., 2002; Kalis et al., 2004) and is recommended for establishment of a MAP-negative herd status in control programs (Collins et al., 2005). Kalis et al. (2004) studied

90 herds with no previous history of paratuberculosis, and which were closed for 3 yr prior to study initiation, as well as during the 5 yr longitudinal study. Pooled fecal cultures were collected every 6 mo from the closed herds. Fourteen percent of herds were detected by pooled fecal culture at the start of the study, and 61% of the herds were found to be infected by the end of the study, indicating that repetition of pooled fecal culture minimizes risk of misclassification of low prevalence MAP-infected herds (Kalis et al., 2004). Our research project was also designed with repeated pooled fecal culture testing to provide a strong herd case definition, against which diagnostic tests were evaluated.

1.10 Thesis Objectives

A case definition is required for diagnostic test evaluation, providing a practical definition of the target condition using a reference standard (Gardner et al., 2011). At the herd-level, testing options for determination of herd MAP status include individual or pooled fecal culture, culture of environmental samples, and ELISA on milk or serum. Performing individual cow fecal cultures on an entire herd is generally cost-prohibitive for determining herd MAP-status, as fecal culture is the most expensive cow-level test (Tiwari et al., 2006). Individual fecal culture specificity approaches 100% (Nielsen and Toft, 2008), but sensitivity is limited, with estimates ranging from 23% (McKenna et al., 2005) to 29% (Whitlock et al., 2000) for infected cows and 74 to 89% for infectious animals, depending on the culture method used (Sockett et al., 1992). Pooling individual cow fecal samples by cow-age, with 5 cows per pooled fecal culture, substantially decreases cost and is both highly sensitive and specific, relative to individual culture, for determination of herd MAP status (Kalis et al., 2004). Wells et al. (2003) reported that pooled fecal culture confirmed 16 of the 17 herds detected with individual fecal culture,

equating to a herd-level sensitivity (**HSe**) of 94% for pooled fecal culture. Nine of the 17 herds had a prevalence of infection < 10%, based on individual fecal culture, and the 1 herd detected positive by individual culture, but negative with pooled fecal culture, had only a single light shedding cow. In the same study, herd-level specificity (**HSp**) of 100% for pooled fecal culture, relative to individual fecal culture, was reported (Wells et al., 2003). Repeated pooled fecal culture was used in this project to establish a herd-level case definition, against which herd-level tests were evaluated.

At the cow-level, case definitions vary based on the target condition of interest. In Chapter 5, the target condition was a MAP-infectious cow, which is a cow excreting adequate amounts of MAP to possibly infect a non-infected cow (Gardner et al., 2011). MAP-infectious cows are a subset of MAP-infected cows. In Chapter 6, the target condition was a MAP-infected cow, which is a cow that has MAP in its tissues (Gardner et al., 2011). For both target conditions, case definitions were developed from results of individual fecal culture.

1.10.1 Herd-level Detection of MAP: Environmental Culture

Environmental culture (**EC**) is a relatively cost-effective and non-invasive herd diagnostic tool for MAP detection (Berghaus et al., 2006; Lombard et al., 2006b). The majority of EC research to-date has been performed in large herds with a history of paratuberculosis or higher within-herd MAP prevalence than what might be expected in many dairy herds. For example, the proportion of MAP-positive herds within EC studies ranged from 53% (Lombard et al., 2006b) to 100% (Berghaus et al., 2006). Fecal culture within-herd prevalence ranged from 2.0% (Pillars et al., 2009a) to 13.6% (Smith et al., 2011). The performance of EC is inferior in low prevalence herds compared to herds with

a higher prevalence (Smith et al., 2009; Pillars et al., 2009a; Pillars et al., 2009b; Smith et al., 2011; Tavornpanich et al., 2012), and there is a need for a more thorough investigation of EC test characteristics in low prevalence herds (Pillars et al., 2009a; Pillars et al., 2009b). Although control programs and consensus documents recommend that an EC sampling protocol includes the collection of 6 composite samples from the herd environment (Collins et al., 2006; USDA, 2008), different sampling procedures were used in previous evaluations of EC (Raizman et al., 2004; Berghaus et al., 2006; Pillars et al., 2009a; Pillars et al., 2009b). The objectives of Chapter 2 are therefore to evaluate EC test characteristics, using an industry-recommended sampling protocol applied to test-negative, low prevalence and high prevalence herds. In this way, the performance of EC in a practical field setting, similar to what would be expected in comparable dairy industries, could be evaluated.

1.10.2 Herd-level Detection of MAP: Milk ELISA

A second method of determining herd MAP status is to test individual lactating cows with a milk ELISA and compile results at the herd-level. Milk ELISA herd testing is both inexpensive and rapid compared to fecal culture (Tiwari et al., 2006), and is easy to implement in dairy herds which participate in Dairy Herd Improvement (**DHI**) programs involving monthly individual cow milk collection. Reported estimates of milk ELISA HSe are 61% (Lombard et al., 2006a), 83% (Klausen et al., 2003), and 92% (Hendrick et al., 2005a). However, these estimates were limited in their applicability because the study herds were not fully representative of the industry (Klausen et al., 2003), or a herd reference standard was not used to evaluate the milk ELISA results (Hendrick et al., 2005a). In addition, the impact of within-herd MAP prevalence on milk

ELISA HSe has not been evaluated (Lombard et al., 2006a). Within-herd prevalence impacts herd-level test characteristics (Christensen and Gardner, 2000), and it is necessary to quantify this effect when attempting to apply the test at the herd-level.

As published in consensus documents, the cow-level specificity of the milk ELISA is slightly lower than that of fecal culture (99.0 ± 1.0 for milk ELISA compared to 99.9 ± 0.1 for fecal culture) (Collins et al., 2006). Although this difference may seem minor, when the test is applied at the herd-level in a MAP-negative or very low prevalence herd, the imperfections in specificity at the cow-level are magnified. Given the potential negative impact that a false-positive herd diagnosis can have, it is imperative to evaluate the HSp of the milk ELISA. Estimates of milk ELISA HSp range from 66% (Hendrick et al., 2005a) to 100% (Klausen et al., 2003). These HSp estimates are relatively unreliable because the studies had either few or no MAP-negative herds included (Klausen et al., 2003), or the herd MAP-status was unknown (Hendrick et al., 2005a). The objectives of Chapter 3 are to evaluate herd-level test characteristics of the milk ELISA, and to investigate the impact of herd-level factors on these test characteristics.

1.10.3 Cow-level Agreement Between Three Commercial Milk ELISAs

Several commercial milk ELISA kits are available for cow-level paratuberculosis diagnosis. Fecal culture is often used as a reference test for evaluation of milk ELISA performance. However, milk ELISA measures antibody response, which can occur more than a year prior to the bacterial shedding detected with FC (Nielsen, 2008). Given the relatively low sensitivity of the milk ELISA relative to fecal culture, it is important to determine if the milk ELISA test results at minimum agree with each other, and also if

false-positive and false-negative results are consistent between kits. The number of studies evaluating agreement between serum ELISAs or between serum and milk ELISA kits is limited (Hendrick et al., 2005b; McKenna et al., 2006a; Lombard et al., 2006a). However, agreement among commercial milk ELISA kits has not been evaluated. The first objective of Chapter 4 is to quantify between-test agreement of commercial milk ELISA kits at the cow-level.

Although milk ELISA Se was found to increase with progressive stages of paratuberculosis (Nielsen and Toft, 2008), and also with cow age (Nielsen and Toft, 2006), it is not known if agreement between milk ELISA tests is similarly affected by these cow-level factors. The influence of within-herd MAP prevalence on between-test agreement is also not known. Consensus recommendations are that the milk ELISA may be an appropriate cow-level test for control in herds with > 10% MAP-positive ELISA results (Collins et al., 2006). However, paratuberculosis is considered a disease of low to moderate within-herd prevalence (Lombard, 2011). As a result, many producers and veterinarians may attempt to interpret cow-level milk ELISA results in herds below the consensus prevalence recommendation. Therefore, the second objective of Chapter 4 is to evaluate herd and cow-level factors influencing positive milk ELISA between-test agreement.

1.10.4 Cow-level Test Characteristics and Likelihood Ratio Interpretation of Commercial Milk ELISA Kits

Estimates of test characteristics in MAP infectious cattle range from 21% (Lombard et al., 2006a) to 61% (Hendrick et al., 2005b) for sensitivity, and from 95% (Klausen et al., 2003; Hendrick et al., 2005b) to 98% (Lombard et al., 2006a) for

specificity. For MAP-infected cattle, sensitivity estimates are sparse, and have been estimated up to 39% (Nielsen et al., 2002). Specificity estimates for detection of MAP-infected cattle range from 96% (Nielsen et al., 2002) to just under 100% (Collins et al., 2005). Comparison between these estimates is limited because of differences in factors such as study population, case definitions, and statistical methods. In addition, many of these milk ELISA test characteristics have been based on in-house procedures that are not commercially available (Nielsen et al., 2002; Klausen et al., 2003; Collins et al., 2005; Hendrick et al., 2005b; Lombard et al., 2006a). Because test characteristics can vary substantially between different kits, it is very important to understand test characteristics of the particular ELISA being applied to a herd (Nielsen, 2009), which in practice is often a commercial ELISA kit.

In addition to limited use of commercially available kits in previous publications, there is also a limited number of milk ELISA specificity estimates from non-infected herds. For paratuberculosis, specificity of milk ELISAs is best estimated within herds that had multiple negative reference tests of all adult cattle (Collins et al., 2005). Collins et al. (2005) evaluated an in-house milk ELISA in 7 uninfected herds (359 cows) and estimated a specificity of just under 100%. There is only 1 published paper estimating a commercial milk ELISA kit's specificity from established non-infected herds (Van Weering et al., 2007). In that study, data were used from 10 herds (435 cows) certified MAP-free in the Dutch control program to evaluate test characteristics of a commercial milk ELISA, and specificity was 100%.

Two statistical methods are commonly used to estimate diagnostic test characteristics. The first, and more traditional approach, is to compare the test under

evaluation to a pseudogold (reference) standard. This approach has been used in several studies estimating milk ELISA test characteristics (Collins et al., 2005; Hendrick et al., 2005b; Lombard et al., 2006a; van Weering et al., 2007), and therefore an estimate produced using a pseudogold standard can be readily compared to other publications. As an alternative, Bayesian latent class models allow evaluation of test characteristics in the absence of a gold standard reference test (Hui and Walter, 1980; Vacek, 1985; Joseph et al., 1995; Enoe et al., 2000; Dendukuri and Joseph, 2001). Given that there is no perfect ante-mortem diagnostic test for paratuberculosis (Gardner et al., 2011), an advantage of this methodology is that the test under evaluation is not compared to a reference test, but rather the latent disease state of the cow. The first objective of Chapter 5 is to evaluate cow-level milk ELISA sensitivity and specificity for 2 commercially available milk ELISA kits to detect MAP-infectious cows, using both pseudogold standard methods and latent class models (no gold or pseudogold standard).

Commercial milk ELISA kits produce a quantitative answer, which is converted to a dichotomous result, based on a manufacturer-recommended cutoff. This dichotomization potentially results in loss of information, as the magnitude of an ELISA response correlates with MAP-infectious status (Nielsen, 2007). Levels of ELISA response represent distinct likelihoods for a cow to be MAP infectious, and the magnitude of this likelihood can be estimated with likelihood ratios (**LR**) (Collins, 2002; Collins et al., 2005), which create a simple decision-making tool for producers and veterinarians, based on quantitative ELISA values (Collins, 2002; Naugle et al., 2003). Collins (2002) pointed out that these LR may be geographically specific. Coupled with a lack of published LRs for commercial kits applied to milk samples, it is useful and

practical to estimate categorical LR, using commercial ELISA kits, which can be applied to the dairy industry. The second objective of Chapter 5 is therefore to evaluate the likelihood of a cow to be MAP-infectious, based on categories of quantitative ELISA output.

1.10.5 Evaluation of Repeated Cow-level Milk ELISA Testing

Given the chronic progressive nature of paratuberculosis, a potential strategy to improve detection of MAP-infected cows is repeated milk ELISA testing. This testing strategy involves both an initial milk ELISA test, as well as a follow-up milk ELISA test, repeated at a specified test interval. Rather than attempting to diagnose a cow using a single ELISA result, initial and repeat test results are evaluated together. In cows > 3 yrs of age, repeated milk ELISA results perform significantly better than a single result (Huda et al., 2004). However, it has also been reported that there is variability in repeated serum ELISA tests (Hirst et al., 2002), and that comparing current serum ELISA output to previous results provides minimal advantage over evaluating the current ELISA result (Sweeney et al., 2006).

Although repeated milk ELISA testing has the potential to increase detection of MAP-infected cows, it is necessary to quantify the information gained from a repeat ELISA test before recommending this practice to producers. If repeat testing has merit, it is also necessary to evaluate a testing interval that would potentially be both economically feasible and convenient for producers. For example, a test interval applied annually at a pre-determined point in a cow's lactation cycle, using a milk sample that is already being collected as part of routine herd milk testing, may be acceptable to many producers from both a convenience and cost perspective. Consideration of cow-level

factors, such as age at repeat testing and magnitude of initial ELISA test, which have both been shown to correlate with progression of paratuberculosis (Nielsen and Toft, 2006; Nielsen, 2007), may be useful strategies to further improve the probability of detecting MAP-infected cows using a repeat milk ELISA.

The first objective of Chapter 6 is to investigate the sensitivity and specificity of initial and repeat milk ELISA combinations to detect MAP-infected cows, at both 6 and 12 mo test intervals. The second objective is to evaluate factors influencing the probability of a MAP infected cow, which was initially milk ELISA negative, to be positive on a repeat ELISA test.

1.10.6 Summary of Objectives

The overall objective of this thesis was to evaluate herd-level test characteristics of EC and milk ELISA, and cow-level test characteristics of milk ELISA, for detection of paratuberculosis. The study was designed within a longitudinal framework to allow for a decreased risk of misclassification of low-prevalence MAP-positive herds, and herds were selected to be representative of the Canadian dairy industry, with respect to within-herd MAP prevalence. Specific objectives were to:

- study herd and cow-level diagnostic tests within a study population of MAP-negative, low and moderate to high within-herd prevalence herds, in order to maximize external validity to the dairy industry;
- evaluate the herd-level test characteristics of EC, and the impact of within-herd MAP prevalence on test characteristics;
- estimate the herd-level test characteristics of commercial milk ELISAs, and the impact of within-herd MAP prevalence on test characteristics;

- describe agreement between 3 commercial milk ELISA kits at the cow-level;
- investigate cow-level test characteristics of commercially available milk ELISA kits;
- develop a categorical interpretation of quantitative milk ELISA results; and
- assess the cow-level diagnostic information gained with a repeated milk ELISA test.

1.11 References

Abbas, B. and H. P. Riemann. 1988. IgG, IgM and IgA in the serum of cattle naturally infected with Mycobacterium paratuberculosis. Comp. Immunol. Microbiol. Infect. Dis. 11:171-175.

Abbas, B., H. Riemann, and D. Hird. 1983. Diagnosis of Johne's disease (paratuberculosis) in northern California cattle and a note on its economic significance. Calif. Vet. 37:20-24.

Abubakar, I., D. Myhill, S. Aliyu, and P. Hunter. 2008. Detection of Mycobacterium avium subspecies paratuberculosis from patients with Crohn's disease using nucleic acid-based techniques: A systematic review and meta-analysis. Inflamm. Bowel Dis. 14:401-410.

Bannantine, J. P. and J. R. Stabel. 2002. Killing of Mycobacterium avium subspecies paratuberculosis within macrophages. BMC Microbiol. 2:2.

Bannantine, J. P., Q. Zhang, L. Li, and V. Kapur. 2003. Genomic homogeneity between Mycobacterium avium subsp. avium and Mycobacterium avium subsp. paratuberculosis belies their divergent growth rates. BMC Microbiol. 3:10.

Barkema, H. W., S. Hendrick, J. deBuck, S. Ghosh, G. G. Kaplan, and K. Rioux. 2011. Crohn's disease in humans and Johne's disease in cattle-Linked diseases? Pages 197-213 in Zoonotic Pathogens in the Food Chain. D. O. Krause and S. Hendrick ed. CAB International, Wallingford, UK.

Begg, D. J. and R. J. Whittington. 2008. Experimental animal infection models for Johne’s disease, an infectious enteropathy caused by Mycobacterium avium subsp. paratuberculosis. Vet. J. 176:129-145.

Benedictus, G., A. Dijkhuizen, and J. Stelwagen. 1987. Economic losses due to paratuberculosis in dairy cattle. Vet. Rec. 121:142-146.

Berghaus, R. D., T. B. Farver, R. J. Anderson, C. C. Jaravata, and I. A. Gardner. 2006. Environmental sampling for detection of Mycobacterium avium ssp. paratuberculosis on large California dairies. J. Dairy Sci. 89:963-970.

Bölske, G. and D. Herthnek. 2010. Diagnosis of paratuberculosis by PCR. Pages 267-283 in Paratuberculosis: Organism, disease, control. M. A. Behr and D. M. Collins ed. CAB International, Wallingford, UK.

Brenner, H. and O. Gefeller. 1997. Variation of sensitivity, specificity, likelihood ratios and predictive values with disease prevalence. Stat. Med. 16:981-991.

Bryant, B., D. Blyde, G. Eamens, and R. Whittington. 2012. Mycobacterium avium subspecies paratuberculosis cultured from the feces of a southern black rhinoceros (Diceros bicornis minor) with diarrhea and weight loss. J. Zoo Wildl. Med. 43:391-393.

Bull, T. J., E. J. McMinn, K. Sidi-Boumedine, A. Skull, D. Durkin, P. Neild, G. Rhodes, R. Pickup, and J. Hermon-Taylor. 2003. Detection and verification of Mycobacterium avium subsp. paratuberculosis in fresh ileocolonic mucosal biopsy specimens from individuals with and without Crohn's disease. J. Clin. Microbiol. 41:2915-2923.

Chi, J., J. A. VanLeeuwen, A. Weersink, and G. P. Keefe. 2002. Direct production losses and treatment costs from bovine viral diarrhoea virus, bovine leukosis virus, Mycobacterium avium subspecies paratuberculosis, and Neospora caninum. Prev. Vet. Med. 55:137-153.

Chiodini, R. J. 1996. Immunology: Resistance to paratuberculosis. Vet. Clin. North Am. Food Anim. Pract. 12:313-343.

Chiodini, R. J., H. J. Van Kruiningen, and R. S. Merkal. 1984. Ruminant paratuberculosis (Johne's disease): The current status and future prospects. Cornell Vet. 74:218-262.

Christensen, J. and I. A. Gardner. 2000. Herd-level interpretation of test results for epidemiologic studies of animal diseases. Prev. Vet. Med. 45:83-106.

Clarke, C. 1997. The pathology and pathogenesis of paratuberculosis in ruminants and other species. J. Comp. Pathol. 116:217-261.

Collins, M. T. 2003. Paratuberculosis: Review of present knowledge. Acta Vet. Scand. 44:217-221.

Collins, M. 2002. Interpretation of a commercial bovine paratuberculosis enzyme-linked immunosorbent assay by using likelihood ratios. Clin. Diagn. Lab. Immunol. 9:1367-1371.

Collins, M. T., I. A. Gardner, F. B. Garry, A. J. Roussel, and S. J. Wells. 2006. Consensus recommendations on diagnostic testing for the detection of paratuberculosis in cattle in the United States. J. Am. Vet. Med. Assoc. 229:1912-1919.

Collins, M. T., S. J. Wells, K. R. Petrini, J. E. Collins, R. D. Schultz, and R. H. Whitlock. 2005. Evaluation of five antibody detection tests for diagnosis of bovine paratuberculosis. Clin. Diagn. Lab. Immunol. 12:685-692.

Cook, K. L., J. S. Britt, and C. H. Bolster. 2010. Survival of Mycobacterium avium subsp. paratuberculosis in biofilms on livestock watering trough materials. Vet. Microbiol. 141:103-109.

Dendukuri, N. and L. Joseph. 2001. Bayesian approaches to modeling the conditional dependence between multiple diagnostic tests. Biometrics. 57:158-167.

Dorshorst, N. C. and J. E. Lombard. 2006. Decision analysis model for paratuberculosis control in commercial dairy herds. Prev. Vet. Med. 75:92-122.

Eisenberg, S., M. Nielen, J. Hoeboer, V. Rutten, D. Heederik, and A. Koets. 2012. Environmental contamination with Mycobacterium avium subspecies paratuberculosis within and around a dairy barn under experimental conditions. J. Dairy Sci.95:6477-6482.

Eisenberg, S. W., A. P. Koets, M. Nielen, D. Heederik, R. Mortier, J. De Buck, and K. Orsel. 2011. Intestinal infection following aerosol challenge of calves with Mycobacterium avium subspecies paratuberculosis. Vet. Res. 42:1-9.

Eisenberg, S. W., M. Nielen, W. Santema, D. J. Houwers, D. Heederik, and A. P. Koets. 2010. Detection of spatial and temporal spread of Mycobacterium avium subsp. paratuberculosis in the environment of a cattle farm through bio-aerosols. Vet. Microbiol. 143:284-292.

Ellingson, J. L., C. A. Bolin, and J. R. Stabel. 1998. Identification of a gene unique to Mycobacterium avium subspecies paratuberculosis and application to diagnosis of paratuberculosis. Mol. Cell. Probes. 12:133-142.

Enoe, C., M. P. Georgiadis, and W. O. Johnson. 2000. Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease state is unknown. Prev. Vet. Med. 45:61-81.

Espejo, L. A., N. Kubat, S. M. Godden, and S. J. Wells. 2013. Effect of delayed exposure of cattle to Mycobacterium avium subsp paratuberculosis on the development of subclinical and clinical Johne's disease. Am. J. Vet. Res. 74:1304-1310.

Fecteau, M. and R. H. Whitlock. 2011. Treatment and chemoprophylaxis for paratuberculosis. Vet. Clin. North Am. Food Anim. Pract. 27:547-557.

Feller, M., K. Huwiler, R. Stephan, E. Altpeter, A. Shang, H. Furrer, G. E. Pfyffer, T. Jemmi, A. Baumgartner, and M. Egger. 2007. Mycobacterium avium subspecies paratuberculosis and Crohn's disease: A systematic review and meta-analysis. Lancet Infect. Dis. 7:607-613.

Gardner, I. A., S. S. Nielsen, R. J. Whittington, M. T. Collins, D. Bakker, B. Harris, S. Sreevatsan, J. E. Lombard, R. Sweeney, D. R. Smith, J. Gavalchin, and S. Eda. 2011. Consensus-based reporting standards for diagnostic test accuracy studies for paratuberculosis in ruminants. Prev. Vet. Med. 101:18-34.

Garry, F. 2011. Control of paratuberculosis in dairy herds. Vet. Clin. North Am. Food Anim. Pract. 27:599-607.

Green, E., M. Tizard, M. Moss, J. Thompson, D. Winterbourne, J. McFadden, and J. Hermon-Taylor. 1989. Sequence and characteristics or IS900, an insertion element identified in a human Crohn's disease isolate or Mycobacterium paratuberculosis. Nucleic Acids Res. 17:9063-9073.

Greiner, M. and I. A. Gardner. 2000. Epidemiologic issues in the validation of veterinary diagnostic tests. Prev. Vet. Med. 45:3-22.

Hagan, W. 1938. Age as a factor in susceptibility to Johne's disease. Cornell Vet. 28:34-40.

Harris, N. B. and R. G. Barletta. 2001. Mycobacterium avium subsp. paratuberculosis in veterinary medicine. Clin. Microbiol. Rev. 14:489-512.

He, Z. and J. De Buck. 2010. Localization of proteins in the cell wall of Mycobacterium avium subsp. paratuberculosis K10 by proteomic analysis. Proteome Sci. 8:21.

Hendrick, S., T. Duffield, K. Leslie, K. Lissemore, M. Archambault, and D. Kelton. 2005a. The prevalence of milk and serum antibodies to Mycobacterium avium subspecies paratuberculosis in dairy herds in Ontario. Can. Vet. J. 46:1126-1129.

Hendrick, S. H., T. E. Duffield, D. E. Kelton, K. E. Leslie, K. D. Lissemore, and M. Archambault. 2005b. Evaluation of enzyme-linked immunosorbent assays performed on milk and serum samples for detection of paratuberculosis in lactating dairy cows. J. Am. Vet. Med. Assoc. 226:424-428.

Hirst, H. L., F. B. Garry, and M. D. Salman. 2002. Assessment of test results when using a commercial enzyme-linked immunosorbent assay for diagnosis of paratuberculosis in repeated samples collected from adult dairy cattle. J. Am. Vet. Med. Assoc. 220:1685-1689.

Hostetter, J., E. Steadham, J. Haynes, T. Bailey, and N. Cheville. 2003. Phagosomal maturation and intracellular survival of Mycobacterium avium subspecies paratuberculosis in J774 cells. Comp. Immunol. Microbiol. Infect. Dis. 26:269-283.

Huda, A., G. Jungersen, and P. Lind. 2004. Longitudinal study of interferon-gamma, serum antibody and milk antibody responses in cattle infected with Mycobacterium avium subsp. paratuberculosis. Vet. Microbiol. 104:43-53.

Hui, S. L. and S. D. Walter. 1980. Estimating the error rates of diagnostic tests. Biometrics. 36:167-171.

Hutchings, M. R., K. Stevenson, A. Greig, R. S. Davidson, G. Marion, and J. Judge. 2010. Infection of non-ruminant wildlife by Mycobacterium avium subsp. paratuberculosis. Pages 188-200 in Paratuberculosis: Organism, disease, control. M. A. Behr and D. M. Collins ed. CAB International, Wallingford, UK.

Johnson-Ifearulundu, Y., J. B. Kaneene, and J. W. Lloyd. 1999. Herd-level economic analysis of the impact of paratuberculosis on dairy herds. J. Am. Vet. Med. Assoc. 214:822-825.

Joseph, L., T. W. Gyorkos, and L. Coupal. 1995. Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard. Am. J. Epidemiol. 141:263-272.

Kabara, E. and P. M. Coussens. 2012. Infection of primary bovine macrophages with Mycobacterium avium subspecies paratuberculosis suppresses host cell apoptosis. Cell. Infect. Microbiol. 3:1-10.

Kalis, C. H., M. T. Collins, H. W. Barkema, and J. W. Hesselink. 2004. Certification of herds as free of Mycobacterium paratuberculosis infection: Actual pooled faecal results versus certification model predictions. Prev. Vet. Med. 65:189-204.

Kennedy, D. and G. Benedictus. 2001. Control of Mycobacterium avium subsp. paratuberculosis infection in agricultural species. Rev. Sci. Tech. Off. Int. Epiz. 20:151-179.

Klausen, J., A. Huda, L. Ekeroth, and P. Ahrens. 2003. Evaluation of serum and milk ELISAs for paratuberculosis in Danish dairy cattle. Prev. Vet. Med. 58:171-178.

Koets, A. P., V. P. Rutten, M. de Boer, D. Bakker, P. Valentin-Weigand, and W. van Eden. 2001. Differential changes in heat shock protein, lipoarabinomannan, and purified protein derivative-specific immunoglobulin G1 and G2 isotype responses during bovine Mycobacterium avium subsp. paratuberculosis infection. Infect. Immun. 69:1492-1498.

Larsen, A. B., A. I. Moyle, and E. M. Himes. 1978. Experimental vaccination of cattle against paratuberculosis (Johne's disease) with killed bacterial vaccines: A controlled field study. Am. J. Vet. Res. 39:65-69.

Larsen, A., R. Merkal, and R. Cutlip. 1975. Age of cattle as related to resistance to infection with Mycobacterium paratuberculosis. Am. J. Vet. Res. 36:255-257.

Lepper, A., C. Wilks, M. Kotiw, J. Whitehead, and K. Swart. 1989. Sequential bacteriological observations in relation to cell-mediated and humoral antibody responses of cattle infected with Mycobacterium paratuberculosis and maintained on normal or high iron intake. Aust. Vet. J. 66:50-55.

Lombard, J. E. 2011. Epidemiology and economics of paratuberculosis. Vet. Clin. North Am. Food Anim. Pract. 27:525-535.

Lombard, J., I. Gardner, S. Jafarzadeh, C. Fossler, B. Harris, R. Capsel, B. Wagner, and W. Johnson. 2012. Herd-level prevalence of Mycobacterium avium subsp. paratuberculosis infection in United States dairy herds in 2007. Prev. Vet. Med. 108:234-238.

Lombard, J., R. Capsel, B. Wagner, J. Payeur, and C. Fossler. 2008. Changes in management practices and herd-level prevalence of Mycobacterium avium subspecies paratuberculosis (MAP) infection on US dairy operations. Proc. Am. Assoc. Bov. Pract. 41:242.

Lombard, J. E., T. M. Byrem, B. A. Wagner and B. J. McCluskey. 2006a. Comparison of milk and serum enzyme-linked immunosorbent assays for diagnosis of Mycobacterium avium subspecies paratuberculosis infection in dairy cattle. J. Vet. Diagn. Invest. 18:448-458.

Lombard, J. E., B. A. Wagner, R. L. Smith, B. J. McCluskey, B. N. Harris, J. B. Payeur, F. B. Garry and M. D. Salman. 2006b. Evaluation of environmental sampling and culture to determine Mycobacterium avium subspecies paratuberculosis distribution and herd infection status on US dairy operations. J. Dairy Sci. 89:4163-4171.

Lombard, J. E., F. B. Garry, B. J. McCluskey and B. A. Wagner. 2005. Risk of removal and effects on milk production associated with paratuberculosis status in dairy cows. J. Am. Vet. Med. Assoc. 227:1975-1981.

McKenna, S. L., H. W. Barkema, G. P. Keefe, and D. C. Sockett. 2006a. Agreement between three ELISAs for Mycobacterium avium subsp. paratuberculosis in dairy cattle. Vet. Microbiol. 114:285-291.

McKenna, S. L., G. P. Keefe, A. Tiwari, J. VanLeeuwen, and H. W. Barkema. 2006b. Johne's disease in Canada part II: Disease impacts, risk factors, and control programs for dairy producers. Can. Vet. J. 47:1089-1099.

McKenna, S. L., G. P. Keefe, H. W. Barkema, and D. C. Sockett. 2005. Evaluation of three ELISAs for Mycobacterium avium subsp. paratuberculosis using tissue and fecal culture as comparison standards. Vet. Microbiol. 110:105-111.

Mikkelsen, H., G. Jungersen, and S. S. Nielsen. 2009. Association between milk antibody and interferon-gamma responses in cattle from Mycobacterium avium subsp. paratuberculosis infected herds. Vet. Immunol. Immunopathol. 127:235-241.

Momotani, E., D. Whipple, A. Thiermann, and N. Cheville. 1988. Role of M cells and macrophages in the entrance of Mycobacterium paratuberculosis into domes of ileal Peyer's patches in calves. Vet. Pathol. 25:131-137.

Mortier, R. A., H. W. Barkema, J. M. Bystrom, O. Illanes, K. Orsel, R. Wolf, G. Atkins, and J. De Buck. 2013. Evaluation of age-dependent susceptibility in calves infected with two doses of Mycobacterium avium subspecies paratuberculosis using pathology and tissue culture. Vet. Res. 44:94.

Münster, P., K. Fechner, I. Völkel, A. von Buchholz, and C. Czerny. 2012. Distribution of Mycobacterium avium ssp. paratuberculosis in a German zoological garden determined by IS900 semi-nested and quantitative real-time PCR. Vet. Microbiol. 163:116-123.

Naser, S. A., G. Ghobrial, C. Romero, and J. F. Valentine. 2004. Culture of Mycobacterium avium subspecies paratuberculosis from the blood of patients with Crohn's disease. Lancet. 364:1039-1044.

Naugle, A. L., W. P. Shulaw, W. J. Saville, T. E. Wittum, and B. Byrum. 2003. Calculation method for likelihood ratios dictates interpretation. Clin. Diagn. Lab. Immunol. 10:729-730.

Nielsen, S. S. 2009. Use of diagnostics for risk-based control of paratuberculosis in dairy herds. In Pract. 31:150-154.

Nielsen, S. S. 2008. Transitions in diagnostic tests used for detection of Mycobacterium avium subsp. paratuberculosis infections in cattle. Vet. Microbiol. 132:274-282.

Nielsen, S. S. 2007. Danish control programme for bovine paratuberculosis. Cattle Pract. 15:161-168.

Nielsen, S. S. and N. Toft. 2008. Ante mortem diagnosis of paratuberculosis: A review of accuracies of ELISA, interferon-gamma assay and faecal culture techniques. Vet. Microbiol. 129:217-235.

Nielsen, S. S. and N. Toft. 2006. Age-specific characteristics of ELISA and fecal culture for purpose-specific testing for paratuberculosis. J. Dairy Sci. 89:569-579.

Nielsen, S. S., M. A. Krogh, and C. Enevoldsen. 2009. Time to the occurrence of a decline in milk production in cows with various paratuberculosis antibody profiles. J. Dairy Sci. 92:149-155.

Nielsen, S. S., C. Gronbaek, J. F. Agger, and H. Houe. 2002. Maximum-likelihood estimation of sensitivity and specificity of ELISAs and faecal culture for diagnosis of paratuberculosis. Prev. Vet. Med. 53:191-204.

Nordlund, K. V., W. J. Goodger, J. Pelletier, and M. T. Collins. 1996. Associations between subclinical paratuberculosis and milk production, milk components, and somatic cell counts in dairy herds. J. Am. Vet. Med. Assoc. 208:1872-1876.

O'Brien, R., C. Mackintosh, D. Bakker, M. Kopecna, I. Pavlik, and J. Griffin. 2006. Immunological and molecular characterization of susceptibility in relationship to bacterial strain differences in Mycobacterium avium subsp. paratuberculosis infection in the red deer (Cervus elaphus). Infect. Immun. 74:3530-3537.

Ontario Johne's Education and Management Assistance Program. 2013. Johne's program stats as of March 23, 2013. Accessed November 18, 2013. http://www.johnes.ca/pdf%20files/Program%20Stats%20as%20of%20Mar%20%2023%202013_xlsx.pdf.

Ott, S. L., S. J. Wells, and B. A. Wagner. 1999. Herd-level economic losses associated with Johne's disease on US dairy operations. Prev. Vet. Med. 40:179-192.

Patton, E. A. 2011. Paratuberculosis vaccination. Vet. Clin. North Am. Food Anim. Pract. 27:573-580.

Pfyffer, G. E. 2006. Mycobacterium: General characteristics, laboratory detection, and staining procedures. Pages 543-572 in Manual of clinical microbiology: Volume 1. P. Murray, E. Baron, J. Jorgensen, M. Landry, and M. Pfaller ed. CAB International, Wallingford, UK.

Pillars, R. B., D. L. Grooms, J. A. Woltanski, and E. Blair. 2009a. Prevalence of Michigan dairy herds infected with Mycobacterium avium subspecies paratuberculosis as determined by environmental sampling. Prev. Vet. Med. 89:191-196.

Pillars, R. B., D. L. Grooms, and J. B. Kaneene. 2009b. Longitudinal study of the distribution of Mycobacterium avium subsp. paratuberculosis in the environment of dairy herds in the Michigan Johne's disease control demonstration herd project. Can. Vet. J. 50:1039-1046.

Raizman, E. A., S. J. Wells, P. A. Jordan, G. D. DelGiudice, and R. R. Bey. 2005. Mycobacterium avium subsp. paratuberculosis from free-ranging deer and rabbits surrounding Minnesota dairy herds. Can. J. Vet. Res. 69:32-38.

Raizman, E. A., S. J. Wells, S. M. Godden, R. F. Bey, M. J. Oakes, D. C. Bentley, and K. E. Olsen. 2004. The distribution of Mycobacterium avium ssp. paratuberculosis in the environment surrounding Minnesota dairy farms. J. Dairy Sci. 87:2959-2966.

Salgado, M., M. Collins, F. Salazar, J. Kruze, G. Bölske, R. Söderlund, R. Juste, I. Sevilla, F. Biet, and F. Troncoso. 2011. Fate of Mycobacterium avium subsp. paratuberculosis after application of contaminated dairy cattle manure to agricultural soils. Appl. Environ. Microbiol. 77:2122-2129.

Sanderson, J., M. Moss, M. Tizard, and J. Hermon-Taylor. 1992. Mycobacterium paratuberculosis DNA in Crohn's disease tissue. Gut. 33:890-896.

Seitz, S. E., L. Heider, W. Hueston, S. Bech-Nielsen, D. Rings, and L. Spangler. 1989. Bovine fetal infection with Mycobacterium paratuberculosis. J. Am. Vet. Med. Assoc. 194:1423-1426.

Singh, S. V., A. V. Singh, A. Kumar, P. K. Singh, R. Deb, A. K. Verma, A. Kumar, R. Tiwari, S. Chakraborty, and K. Dhama. 2013. Survival mechanisms of Mycobacterium avium subspecies paratuberculosis within host species and in the environment-A review. Nat. Sci. 5:710-723.

Smith, R. L., Y. H. Schukken, A. K. Pradhan, J. M. Smith, R. H. Whitlock, Van Kessel J.S., D. R. Wolfgang, and Y. T. Grohn. 2011. Environmental contamination with Mycobacterium avium subsp. paratuberculosis in endemically infected dairy herds. Prev. Vet. Med. 102:1-9.

Smith, R. L., Y. T. Grohn, A. K. Pradhan, R. H. Whitlock, J. S. Van Kessel, J. M. Smith, D. R. Wolfgang, and Y. H. Schukken. 2009. A longitudinal study on the impact of Johne's disease status on milk production in individual cows. J. Dairy Sci. 92:2653-2661.

Sockett, D. C., D. J. Carr, and M. T. Collins. 1992. Evaluation of conventional and radiometric fecal culture and a commercial DNA probe for diagnosis of Mycobacterium paratuberculosis infections in cattle. Can. J. Vet. Res. 56:148-153.

Stabel, J. 2000. Transitions in immune responses to Mycobacterium paratuberculosis. Vet. Microbiol. 77:465-473.

Stabel, J. 1998. Johne's disease: A hidden threat. J. Dairy Sci. 81:283-288.

Streeter, R. N., G. Hoffsis, S. Bech-Nielsen, W. Shulaw, and D. Rings. 1995. Isolation of Mycobacterium paratuberculosis from colostrum and milk of subclinically infected cows. Am. J. Vet. Res. 56:1322-1324.

Sweeney, R. W. 2011. Pathogenesis of paratuberculosis. Vet. Clin. North Am. Food Anim. Pract. 27:537-546.

Sweeney, R. W., R. H. Whitlock, S. McAdams, and T. Fyock. 2006. Longitudinal study of ELISA seroreactivity to Mycobacterium avium subsp. paratuberculosis in infected cattle and culture-negative herd mates. J. Vet. Diagn. Invest. 18:2-6.

Sweeney, R. W., R. H. Whitlock, and A. E. Rosenberger. 1992a. Mycobacterium paratuberculosis cultured from milk and supramammary lymph nodes of infected asymptomatic cows. J. Clin. Microbiol. 30:166-171.

Sweeney, R., R. Whitlock, and A. Rosenberger. 1992b. Mycobacterium paratuberculosis isolated from fetuses of infected cows not manifesting signs of the disease. Am. J. Vet. Res. 53:477-480.

Tavornpanich, S., S. J. Wells, C. P. Fossler, A. J. Roussel, and I. A. Gardner. 2012. Evaluation of an alternative method of herd classification for infection with paratuberculosis in cattle herds in the United States. Am. J. Vet. Res. 73:248-256.

Taylor, A. W. 1953. Experimental Johne's disease in cattle. J. Comp. Path. 63:355-367.

Tiwari, A., J. A. VanLeeuwen, I. R. Dohoo, G. P. Keefe, and A. Weersink. 2008. Estimate of the direct production losses in Canadian dairy herds with subclinical Mycobacterium avium subspecies paratuberculosis infection. Can. Vet. J. 49:569-576.

Tiwari, A., J. A. Vanleeuwen, I. R. Dohoo, G. P. Keefe, J. P. Haddad, R. Tremblay, H. M. Scott and T. Whiting. 2007. Production effects of pathogens causing bovine leukosis, bovine viral diarrhea, paratuberculosis, and neosporosis. J. Dairy Sci. 90:659-669.

Tiwari, A., J. A. VanLeeuwen, S. L. McKenna, G. P. Keefe, and H. W. Barkema. 2006. Johne's disease in Canada part I: Clinical symptoms, pathophysiology, diagnosis, and prevalence in dairy herds. Can. Vet. J. 47:874-882.

USDA. 2008. Johne's disease on US dairies, 1991-2007. USDA-APHIS-VS, CEAH, National Animal Health Monitoring System, Fort Collins, CO, USA (#N521.0408).

Vacek, P. M. 1985. The effect of conditional dependence on the evaluation of diagnostic tests. Biometrics. 41:959-968.

van Weering, H., G. van Schaik, A. van der Meulen, M. Waal, P. Franken, and K. van Maanen. 2007. Diagnostic performance of the Pourquier ELISA for detection of antibodies against Mycobacterium avium subspecies paratuberculosis in individual milk and bulk milk samples of dairy herds. Vet. Microbiol. 125:49-58.

VanLeeuwen, J. A., G. P. Keefe, R. Tremblay, C. Power, and J. J. Wichtel. 2001. Seroprevalence of infection with Mycobacterium avium subspecies paratuberculosis, bovine leukemia virus, and bovine viral diarrhea virus in Maritime Canada dairy cattle. Can. Vet. J. 42:193-198.

Wells, S. J., S. M. Godden, C. J. Lindeman, and J. E. Collins. 2003. Evaluation of bacteriologic culture of individual and pooled fecal samples for detection of Mycobacterium paratuberculosis in dairy cattle herds. J. Am. Vet. Med. Assoc. 223:1022-1025.

Wells, S. J., R. H. Whitlock, B. A. Wagner, J. Collins, F. Garry, H. Hirst, J. Lawrence, W. J. Saville, and A. L. Naugle. 2002. Sensitivity of test strategies used in the voluntary Johne's disease herd status program for detection of Mycobacterium paratuberculosis infection in dairy cattle herds. J. Am. Vet. Med. Assoc. 220:1053-1057.

Wells, S. J. and B. A. Wagner. 2000. Herd-level risk factors for infection with Mycobacterium paratuberculosis in US dairies and association between familiarity of the herd manager with the disease or prior diagnosis of the disease in that herd and use of preventive measures. J. Am. Vet. Med. Assoc. 216:1450-1457.

Whitlock, R. H. 2010. Paratuberculosis control measures in the USA. Pages 319-329 in Paratuberculosis: Organism, disease, control. M. A. Behr and D. M. Collins ed. CAB International, Wallingford, UK.

Whitlock, R. H. and C. Buergelt. 1996. Preclinical and clinical manifestations of paratuberculosis (including pathology). Vet. Clin. North Am. Food Anim. Pract. 12:345-356.

Whitlock, R. H., S. J. Wells, R. W. Sweeney, and J. Van Tiem. 2000. ELISA and fecal culture for paratuberculosis (Johne's disease): Sensitivity and specificity of each method. Vet. Microbiol. 77:387-398.

Whittington, R. J. and P. A. Windsor. 2009. In utero infection of cattle with Mycobacterium avium subsp. paratuberculosis: A critical review and meta-analysis. Vet. J. 179:60-69.

Whittington, R. J., I. B. Marsh, and L. A. Reddacliff. 2005. Survival of Mycobacterium avium subsp. paratuberculosis in dam water and sediment. Appl. Environ. Microbiol. 71:5304-5308.

Whittington, R. J., D. J. Marshall, P. J. Nicholls, I. B. Marsh, and L. A. Reddacliff. 2004. Survival and dormancy of Mycobacterium avium subsp. paratuberculosis in the environment. Appl. Environ. Microbiol. 70:2989-3004.

Wilson, D. J., C. Rossiter, H. Han, and P. Sears. 1995. Financial effects of Mycobacterium paratuberculosis on mastitis, milk production, and cull rate in clinically normal cows. Agri-Practice 1995. 16:12-18.

Wilson, D. J., C. Rossiter, H. Han, and P. Sears. 1993. Association of Mycobacterium paratuberculosis infection with reduced mastitis, but with decreased milk production and increased cull rate in clinically normal dairy cows. Am. J. Vet. Res. 54:1851-1857.

Windsor, P. A. and R. J. Whittington. 2010. Evidence for age susceptibility of cattle to Johne’s disease. Vet. J. 184:37-44.

CHAPTER 2

EVALUATION OF ENVIRONMENTAL FECAL CULTURE FOR MYCOBACTERIUM AVIUM SUBSPECIES PARATUBERCULOSIS DETECTION IN DAIRY HERDS AND ASSOCIATION WITH APPARENT WITHIN-HERD PREVALENCE

Carrie J. Lavers, Shawn L. B. McKenna, Ian R. Dohoo, Herman W. Barkema, Greg P. Keefe

The primary contents of this chapter are published as: C. J. Lavers, S. L. B. McKenna, I. R. Dohoo, H. W. Barkema, and G. P. Keefe. 2013. Evaluation of environmental fecal culture for Mycobacterium avium subspecies paratuberculosis detection in dairy herds and association with apparent within-herd prevalence. Can. Vet. J. 54: 1053-1060.

2.1 Abstract

This study evaluated test characteristics of environmental culture (EC) for the detection of Mycobacterium avium subspecies paratuberculosis (MAP) in 32 herds over a 2 yr period. Individual fecal samples were collected every 6 mo and environmental fecal samples every 3 mo. Pooled fecal culture was conducted initially, with subsequent individual fecal culture performed on all lactating cow samples from positive pools. Samples were cultured in a broth culture system (ESP® Culture System II, TREK Diagnostic Systems, Inc., Cleveland, Ohio, USA), with confirmatory PCR (VetAlert™ Johne's Real-Time PCR kit, Tetracore, Inc., Rockville, Maryland, USA) performed on any positive fecal samples. Relative to a MAP herd-status based on all pooled fecal culture results collected during the study, sensitivity (Se) of a set of 6 EC-samples collected from prescribed locations within the herd environment (EC-6) was 71% (95% CI: 49-86%) and specificity (Sp) was 99% (95% CI: 95-100%). Sensitivity of EC increased as apparent within-herd fecal culture prevalence (WHP) increased. As examples, Se increased from 51% (95% CI: 36-66%), when WHP was 2%, to 99% (95% CI: 91-100%) when WHP was 8%. The estimated WHP increased as the proportion of positive EC-samples within an EC-6 set increased. As examples, WHP increased from 0.7% (95% CI: 0.4-1.2%) when 1 of 6 EC-samples was positive, to 12.7% (95% CI: 10.8-14.9%) when all 6 EC-samples were positive. Environmental culture is an acceptable tool for herd diagnosis of MAP in low-prevalence herds, with Se increasing as WHP increases.

2.2 Introduction

Johne's disease (**JD**) is a chronic infectious enteritis of ruminants caused by Mycobacterium avium subspecies paratuberculosis (**MAP**). Control of JD is motivated by 2 factors: the production-limiting effects of the disease, most notably decreased milk production (Nielsen et al., 2009) and premature culling (Lombard et al., 2005) with reduced slaughter weight (Johnson-Ifearulundu et al., 1999), and the concern of a possible association between MAP and Crohn's disease in humans (Barkema et al., 2011).

Determining whether a herd is MAP-positive or MAP negative can be challenging. Performing individual cow fecal cultures on an entire herd in order to establish a herd diagnosis for MAP is generally considered cost-prohibitive, with each cow fecal culture costing approximately $40 CDN (Tiwari et al., 2006). Pooling of individual cow fecal samples (**PFC**) offers a substantial cost-savings, and has a herd-level sensitivity (**Se**) of 94% and herd-level specificity (**Sp**) of 100%, relative to individual fecal culture (Wells et al., 2003). Another herd-level MAP diagnostic tool is environmental culture (**EC**), which is relatively cost-effective and non-invasive (Raizman et al., 2004; Berghaus et al., 2006).

Previous evaluation of the relationship between EC test characteristics and within-herd MAP prevalence has been limited. Pillars et al. (2009) reported within-herd prevalence levels in 7 study herds ranged from 0-42%, but the focus of the research was the distribution of MAP in the environment. Smith et al. (2011) studied 3 herds, which are described by Pradhan and Schukken (2009) as having mean fecal culture within-herd prevalence estimates of 1.5, 2.5, and 5.4%, and cautioned against the use of EC in low-

prevalence herds, as Se was estimated to be 40%. This was a small number of herds from which to draw this conclusion and further investigation is warranted.

It is important to estimate EC test characteristics in low-prevalence herds because MAP is generally considered a disease of low within-herd prevalence (Lombard, 2011). It was anticipated that EC test characteristics would be lower in dairy herds with low MAP within-herd prevalence, relative to EC test characteristics previously published where within-herd prevalence of study herds was not considered in the analyses. The current study was designed within a longitudinal study frame to allow for a decreased risk of misclassification of low-prevalence MAP-positive herds, and therefore a strengthened gold standard herd classification was used to evaluate EC test results. The objectives of the present study were to evaluate the test characteristics of EC within MAP-negative and MAP-positive herds, with the specific focus of evaluating EC test characteristics in herds with low MAP prevalence within purposively selected herds.

2.3 Materials and Methods

2.3.1 Study Population

A total of 34 herds from the 3 Canadian Maritime provinces participated in this 2 yr prospective study. Twenty-seven herds were originally selected to participate based on a non-random selection process. The herds were selected based upon risk assessments completed as part of a previous MAP awareness project at the Atlantic Veterinary College, with the aim to obtain a mixture of MAP-positive low-prevalence, MAP-positive high-prevalence and MAP-negative herds.

After the first round of sampling, there were too few MAP-positive herds, therefore 7 herds were added for the remaining 18 mo of the project. These herds were chosen based on historic MAP-positivity, and all 7 herds were PFC-positive during the study.

2.3.2 Individual Cow and Environmental Manure Sample Collection

In order to establish a reference MAP herd-status, pooled fecal samples were created by cow age from individual fecal samples collected at 6 mo intervals from all lactating cows, using a clean rectal sleeve lubricated with water. Approximately 30 g of feces was collected per rectum and placed in a clean, labeled 95 mL plastic specimen jar. A set of environmental samples, consisting of 6 manure samples collected from specific sites within the herd environment, were collected every 3 mo. This set of 6 samples is referred to as an EC-6 (set), and each of the 6 samples comprising the set is referred to as an EC-sample. The EC-6 samples were collected following a protocol based on the Voluntary Bovine Johne's Disease Control Program (USDA, 2010). Two EC-samples were collected from the manure storage areas (pits, lagoons, manure piles or manure spreaders). Two EC-samples were collected from the mature cow manure concentration areas (alleyways, gutters, adjacent to waterers or feeders). Two EC-samples were collected from mature cow maternity or sick pens if there were 2 or more animals in the pen, and if manure clean-out did not occur between animals. In the majority of farms participating, only 1 cow was in the sick/maternity pens at 1 time, with manure clean-out between cows. In these cases, 2 additional EC-samples (4 EC-samples total) were collected from the manure concentration areas. In tie-stall barns, manure concentration EC-samples were collected from corners and crevices of the gutter and along the paddles

of the stable-cleaner. To create an EC-sample for both tie-stall and free-stall sites, four 'grabs' of manure within the specified collection site were collected into a clean 95 mL specimen cup to form 1 composite EC-sample. Each EC-sample was collected using a clean latex glove. For individual cow and environmental manure samples, samples were kept cool during transport back to the laboratory. If samples could not be processed immediately, they were frozen. Samples were frozen at -20˚C if processed within 2 wk and were frozen at -80˚C if processed between 2 and 12 wk after collection.

2.3.3 Laboratory Testing

All manure samples were processed by the technical laboratory staff at the Maritime Quality Milk Laboratory (Atlantic Veterinary College, Charlottetown, PEI, Canada), which was approved by USDA proficiency-testing for this technique. Individual cow fecal samples were pooled by age into PFC samples, with 5 cows in each PFC. The individual cow samples comprising positive-PFC samples were thawed and cultured individually (**PFC-IC**). Fecal culture procedures have been described previously (Lavers et al., 2013). Briefly, fecal samples were processed and inoculated into ESP para-JEM® broth (Nova Century Scientific, Inc., Burlington, Ontario, Canada), according to the manufacturer's protocol, with the exception that samples were incubated for 49 d, rather than 42 d (Tavornpanich et al., 2004; Kim et al., 2004). The broth bottle was placed in the ESP® Culture System II (TREK Diagnostic Systems, Inc., Cleveland, Ohio, USA) for incubation up to a maximum of 49 d.

Confirmatory tests were initiated when the culture system indicated a sample to be positive by headspace pressure change. If confirmatory PCR was negative, the sample was returned to the culture system. After the 49 d culture was complete, all broth samples

were examined microscopically for the presence of Mycobacteria using an acid-fast stain. A sample from an agitated para-JEM® broth jar was smeared onto a slide, air-dried and heat-fixed. Smears were stained with carbolfuchsin and counter-stained with methylene blue before being examined for the presence of Mycobacteria. As in McKenna et al. (McKenna et al., 2005), final confirmatory PCR was performed on all samples positive by the culture system and/or microscopic visualization. The confirmatory PCR kit, VetAlert™ Johne's Real-Time PCR kit (Tetracore, Inc., Rockville, Maryland, USA), which detected the hspX gene, was performed following manufacturer's instructions. All fecal sample cultures (PFC, PFC-IC and EC-samples) were classified as a positive sample (PFC positive, PFC-IC positive and EC-sample positive, respectively) if both liquid broth culture and confirmatory PCR testing were positive.

Terminology used to describe study parameters is outlined in Table 2.1. The MAP herd-status used for the purpose of evaluating environmental culture was determined from the repeated PFC test results. A MAP test-negative herd was defined as a herd in which no positive-PFC from the herd were detected during the study. A herd was classified as MAP-positive if a positive-PFC was collected during the study, indicating that MAP infectious cows were present in the herd. Because PFC is an imperfect gold standard test, three or more rounds of cow fecal cultures had to be collected from a herd in order to establish MAP herd-status and be included in the analyses. For the purpose of EC test evaluation, it was assumed that the MAP herd-status was constant during the study period. The individual cow classifications were used to establish an apparent within-herd prevalence estimate (**WHP**). Using the mean WHP, a herd was classified as low-prevalence if $< 5\%$ of cows within a herd were PFC-IC positive. If $\geq 5\%$ of cows

within a herd were PFC-IC positive, the herd was classified as high-prevalence. An EC-6 set was considered a positive EC-6 (set) if there were ≥ 1 EC-samples positive within the EC-6 set. An EC-6 set was considered a negative EC-6 (set) if there were no EC-samples positive within the EC-6 set.

2.3.4 Statistical Analyses

Descriptive statistics (percentages, means, medians, ranges) were calculated to describe the sampled herds and test results. Due to the repeated samples of animals and herds, simple Se and Sp calculations would underestimate the width of the 95% confidence interval around these estimates. Therefore, Se and Sp of an EC-6 set were calculated using generalized estimating equations (**GEE**) with an exchangeable correlation structure to account for the repeated measures data.

The estimate of Se was calculated using only the results of the EC-6 sets for MAP-positive herds. A null logistic GEE model, using the model outcome of EC-6 results, determined the probability of detecting an infected herd using the equation: $Se=e^{\mu+}/(1+e^{\mu+})$, where $\mu+=\beta_0+\sum\beta_j X_j$ was the linear predictor from the model (Dohoo et al., 2009). Similarly, Sp of an EC-6 set, relative to a MAP-negative herd-status, was derived from the linear predictor of a null logistic GEE model, using the equation $Sp=1-(e^{\mu-}/(1+e^{\mu-}))$. This methodology, described by Dohoo et al. (2009), allows for incorporation of potentially influential factors. For Se and Sp calculation just described, null models were used, meaning no predictors were included in the models. In calculations of Se and Sp, μ+ and μ- were the intercepts of the respective null models. In subsequent analyses (see below), this method was used to determine predictors influencing Se and Sp.

In order to evaluate predictors influencing Se and Sp of an EC-6 set, logistic GEE models were created. As an example, the Se model was:

$$\text{logit}(p) = \beta_0 + \text{SEASON} + \text{BTYPE} + \text{MAP-prevalence}$$

where p is the probability of a MAP-positive herd to have a positive EC-6 set test result; β_0 is the common intercept; SEASON is a categorical variable indicating season of environmental manure sample collection; BTYPE is the housing facility (free-stall or tie-stall); and MAP-prevalence is the mean FC within-herd MAP prevalence (continuous). If the P-value of these predictors on univariable analysis was ≤ 0.15 they were entered into a multivariable logistic GEE model. Lowess smoothers were generated to evaluate linearity and fractional polynomial models were created to explore power transformations of significant continuous predictors in an effort to optimize linearity.

In order to evaluate if an EC-sample positive result was more likely from a specific location within the herd environment, a multi-level logistic regression model was utilized. This model accounted for clustering of EC-samples within herds, with EC-sample result as the outcome and EC-sample collection location as a predictor.

The impact of repeated EC sampling on EC-6 Se and Sp was evaluated using a pair-wise combination of consecutive collected EC-6 results. Pairs were interpreted as positive if either of the two consecutive EC-6 sets were positive. For this analysis, 1 herd could have up to 7 consecutive pair-wise observations if they had 8 EC samples taken over the 2 yr sampling period. This pair-wise interpretation was used as the outcome in null logistic GEE models, and Se and Sp were calculated as above. Generalized

estimating equation models were repeated with triplicate and upward combinations of EC-6 results (to a maximum of 8 EC-6 results), where a combination of EC-6 results was considered positive if at least 1 EC-6 set in the combination had been EC-6 positive.

A cow-level GEE logistic model was constructed to evaluate if the proportion of positive EC-samples within two consecutive EC-6 sets was predictive of the WHP. Because the outcome variable (PFC-IC result) was based on semi-annual collection, the predictor (proportion positive EC-samples) was combined from quarterly into semi-annual measures. The model structure was:

$$\text{logit}(p) = \beta_0 + \text{PROP\_POS} + \text{SEASON} + \text{HSIZE}$$

where p is the probability for a cow within a herd to be PFC-IC positive; β_0 is the common intercept; PROP_POS is the proportion of positive EC-samples within two consecutive EC-6 sets; SEASON is a categorical variable indicating season of environmental manure sample collection; and HSIZE is the herd size (continuous). If the P-value of these predictors on univariable analyses was ≤ 0.15, they were entered into a multivariable logistic GEE model. Lowess smoothers were generated to evaluate linearity and fractional polynomial models were created to explore power transformations of significant predictors in an effort to optimize linearity. Because the outcome was at the cow-level and the predictors at the herd-level, the outcome of this cow-level model would become the probability for a cow within a herd to be PFC-IC positive. This individual cow-level probability applied to each cow within the herd, and would therefore be analogous to the WHP.

All descriptive and statistical data analyses were conducted using Stata/IC® Version 11.2 (StataCorp LP, College Station, Texas, USA). A P-value ≤ 0.05 was considered significant.

2.4 Results

2.4.1 Herd Demographics

In total, 34 herds participated in the project between April 2009 and March 2011. Two herds did not meet criteria for inclusion in the analyses because of incomplete sampling and were excluded from the analyses. The sampling schedule for 7 herds consisted of 3 individual cow fecal and 6 environmental collections, whereas 25 herds had 4 cow fecal and 8 environmental collections. Median herd size was 66 milking cows (mean: 82; range: 28 to 220). Median herd cow age at testing was 4.1 yrs (mean: 4.1; range: 2.9 to 5.5). Eleven facilities (34%) were tie-stall, and 21 facilities (66%) were free-stall. One herd expanded from 190 to 220 milking cows during the project. This herd had a PFC-positive test prior to the introduction of new animals. Remaining herds, while not necessarily closed, did not have substantial introductions during the project. No herds had sheep or goats on the premises.

2.4.2 Herd Prevalence

Overall herd-status, based on PFC-results from the study period, was MAP-negative for 18 herds (56%) and MAP-positive for 14 herds (44%). Nine MAP-positive herds had ≥ 1 PFC-positive result at each round of sampling, and 5 MAP-positive herds

fluctuated between test-negative PFC results and ≥ 1 PFC-positive result at each herd visit (Table 2.2). For the 5 MAP-positive herds that did not have ≥ 1 PFC-positive result at every round of sampling, 2 herds had 3 of 4 herd PFC collections with ≥ 1 positive PFC result. The remaining 3 MAP-positive herds that did not have ≥ 1 PFC-positive result at each round of sampling had positive PFC results in the second and third rounds of herd PFC collections. Mean WHP for all 32 herds, based on PFC-IC results, ranged from 0 to 15.6%, with a relatively even spread of WHP values between the minimum and maximum WHP (Fig. 2.1). In the case of a positive PFC with negative PFC-IC results, for determination of WHP it was considered that 1 cow in the positive-PFC was MAP-positive. The mean WHP of the 3 herds in which this situation occurred (happening a total of 4 times) was 0.9%.

2.4.3 Environmental Cultures

Summary information of PFC and EC results in herds with ≥ 1 positive PFC or ≥ 1 positive EC-6 is displayed in Table 2.2. One MAP-negative herd (no positive-PFC) had 1 positive EC-6 set and 7 test-negative EC-6 sets (Herd #15). One MAP-positive herd had no positive EC-6 sets collected throughout the study (Herd #14). The MAP-positive herd-status for this herd was the result of 1 positive PFC, from which no positive PFC-IC cows were identified, and the herd's mean WHP was 0.2%. Seven of the 14 MAP-positive herds had ≥ 1 test-negative EC-6 set. The mean WHP for these 7 herds was 2.1%. Twelve of 14 herds with ≥ 1 positive EC-6 set were EC-6 positive in the first round of testing, 1 herd was EC-6 positive in the second round and 1 in the seventh round of EC-6 sampling.

Over the course of the project, there was 1 EC-6 positive set and 138 test-negative EC-6 sets collected from the 18 MAP-negative herds. In the 14 herds that were classified as MAP-positive, over the course of the project, there were 67 positive EC-6 sets and 29 test-negative EC-6 sets collected. In total, there were 235 EC-6 sets collected, and 231 EC-6 sets (1,386 EC-samples) had collection location identified. On average, 20% of the EC-samples from these rounds were positive (Table 2.3). Numerically, MAP was recovered less than 40% as frequently from cow concentration areas (e.g. maternity and sick cow pens) compared to manure concentration (e.g. alleyways or gutters) or manure storage areas. However, when herd was accounted for within a logistic model, collection location was not significantly associated with the EC-sample result ($P = 0.13$).

2.4.4 Test Characteristics of Environmental Culture

Based on null logistic GEE models (Table 2.4), the Se of an EC-6 set was 71% (95% CI: 49-86%) and Sp was 99% (95% CI: 95-100%).

A multivariable logistic GEE model evaluating predictors influencing EC-6 Se demonstrated that sensitivity increased with increasing WHP (Table 2.4 and Fig. 2.2). Being MAP-positive overall did not preclude the herd from having WHP values of 0% at some point(s) throughout the study. Sensitivity was not affected by the season of EC-6 set collection.

When the results of successive EC-6 collections were evaluated for accuracy of determining MAP herd-status, there was no substantial effect of using more than 1 EC-6 set on either the Se or Sp of an EC-6 set, regardless of the number of successive EC-6 sets that were used to classify herds. However, a trend toward improved Se was observed

(Table 2.5) with each successive EC-6 set added. The greatest numerical increase in Se occurred between a single EC-6 set and 2 consecutive EC-6 sets (71% versus 81%) (Fig. 2.3).

2.4.5 Proportion of Positive EC-Samples Within Positive EC-6 Sets

Sixty-eight of the 235 EC-6 sets collected (29%) were positive. In 43% of the EC-6 positive sets, all 6 EC-samples within the set were positive. The mean percentage of positive EC-samples within a positive EC-6 set was 71%, meaning a positive EC-6 set contained, on average, 4 to 5 positive EC-samples. The proportion of positive EC-samples within an EC-6 set was a significant predictor of the probability for a cow within a herd to be PFC-IC positive. This was evaluated in a data set representing repeated testing of 4378 cows from the 32 herds. This cow-level probability is equal for each cow in the herd and is analogous to the WHP. As the proportion of positive EC-samples increased, this within-herd cow-level probability (WHP) increased (Fig. 2.4). For best fit of the model to evaluate if the proportion of positive EC-samples within two consecutive EC-6 sets was predictive of the WHP, the cow-level GEE logistic model required a log transformation of the WHP predictor. No other herd level predictors were found to be significant.

2.5 Discussion

Distinguishing a MAP-negative herd from a MAP-positive, low-prevalence herd can be difficult (Tavornpanich et al., 2012). Test characteristics of EC in MAP-positive, low-prevalence herds have previously been questioned (Smith et al., 2011), and since the

majority of MAP-positive dairy herds are expected to be low-prevalence (Barkema et al., 2010), it is critical to understand how EC will perform in these herds. The wide range of MAP WHP in this study (Table 2.2), and in particular the high proportion of low-prevalence herds, fills a knowledge gap regarding EC test characteristics.

Repeated sampling will maximize the identification of MAP-positive herds (Kalis et al., 2004) and thereby minimize the influence of misclassification bias on herd Se and Sp (Christensen and Gardner, 2000). In the current study, 5 of the 14 MAP-positive herds had at least 1 herd test with no positive PFC samples (Table 2.2). The mean WHP in these 5 herds was 0.8%. As a comparison, the mean WHP in the 9 MAP-positive herds that had ≥ 1 positive PFC at every herd test was 9.1%. These results also indicate that low-prevalence herds are more susceptible to misclassification of MAP herd-status (false-negative herd classification) based on cross-sectional sampling at 1 point in time.

Repeated sampling, in the form of multiple negative tests from all adult cattle in the herd, is required for establishment of a MAP-negative herd status (Collins et al., 2005). All 18 study herds classified as MAP-negative had 4 whole-herd PFC collections with no positive-PFC results, and while they were not designated as closed herds, none of the MAP-negative herds increased in size during the study, making it unlikely that they introduced new animals to the herd that were MAP-positive. It was assumed in this analysis that MAP herd-status was stable over the study period. There is limited field data supporting the short-term efficacy of MAP eradication programs, even in MAP-positive low-prevalence herds (Collins et al., 2006). As a result, a change from MAP-positive to MAP-negative herd status was also considered unlikely in the 2 yr study time frame. Nine of the 14 MAP-positive herds were PFC-positive at all herd collections and the 5

herds with intermittent MAP PFC-positivity were not more likely to be positive at the beginning or end of the 2 yr study window.

In general, studies involving EC focus on MAP-positive herds (Aly et al., 2009; Pillars et al., 2009; Smith et al., 2011), and the Se of EC has been examined more frequently than Sp. Raizman et al. (2004) studied 108 herds, 28 of which had been historically classified as uninfected. An EC sampling program, from which 2 individual environmental samples were collected, resulted in 1 of these 28 herds being classified positive by EC (Sp of 96%). Lombard et al. (2006) sampled 98 herds, and 10 of the 60 herds tested with cow fecal culture were considered negative based on a single fecal culture sampling. Of these 10 herds, 2 herds were EC-positive (Sp of 80%). The 18 herds classified as MAP-negative from PFC in our study are a good subset from which to calculate EC Sp due to the repeated PFC sampling. However, Sp estimates from the current research (Table 2.5) may also be underestimated, as individual cow fecal samples were only collected from the milking herd. Dry cows, which typically represent 15% of the adult cow population in a herd, were not included in the individual cow sampling in order to facilitate sample collection. Therefore, the cows represented in the EC samples were not identical to the individual cows sampled, and therefore the 1 herd with a positive EC-6 test but no PFC-positive tests (Table 2.2) may actually have been classified a MAP-positive herd if all cows had been tested with the PFC at each sampling, possibly eliminating the 1 apparent false-positive EC result. Repeated sampling did increase the opportunity for all cows to be included in a PFC collected during the study, but a MAP-positive dry cow could have been culled peri-parturiently prior to the next PFC of milk cows.

Biologically, EC would be expected to have a Sp of 100%. In this project, a positive culture was confirmed with PCR analysis to ensure the positive culture was not another Mycobacterium species. In this study, there was 1 MAP-negative herd with 1 positive EC-sample within an EC-6 set. The other 7 EC-6 sets collected from this herd were EC-6 test-negative. A low within-herd prevalence is 1 of the most likely reasons a herd would be EC-positive yet MAP-negative.

Previous studies using fecal culture to define MAP herd-status reported Se values for EC ranging from 40 to 74% (Berghaus et al., 2006; Lombard et al., 2006; Aly et al., 2009; Pillars et al., 2009; Smith et al., 2011). Variation in reported Se is due, in part, to the herd in which the test is applied, and Se is expected to be higher in high-prevalence herds (Smith et al., 2011). This was consistent with our study results, and although EC Se was relatively low at very low WHP values (≤ 2% WHP), it increased quickly and approached 100% Se at moderate WHP levels of approximately 8% (Figure 2.2). The WHP estimates may be an underestimation of the true within-herd prevalence for these herds, as PFC-IC were performed only on cows from positive-PFC. A single cow shedding low levels of MAP may not have been detected in the PFC, and therefore would have been a false-negative result, contributing to an underestimation of the true within-herd prevalence. Individual PFC-IC samples also underwent an additional freeze-thaw cycle, which could have contributed to an underestimation of the WHP.

In agreement with previous studies (Lombard et al., 2006), no other herd predictors were significantly associated with Se or Sp. Only 1 MAP-positive herd had tie-stall facilities, precluding statistical evaluation of the relationship between housing type and EC Se. In this 50-cow tie-stall herd, 6 of 7 EC-6 sets collected were positive, and

mean WHP was 2.9%. Although not a statistical evaluation, this suggests it is possible to detect MAP-positive, low-prevalence herds in tie-stall facilities using EC. Further study into the Se of EC for tie-stall facilities is required.

When evaluating collection location in all study herds, manure storage and concentration areas had a numerically greater, although not statistically significant, number of positive EC-samples than cow concentration areas (Table 2.3). Previous studies have also reported manure storage and shared alleyways as the sites most likely to be positive (Raizman et al., 2004; Pillars et al., 2009; Smith et al., 2011). Based on current and previous study results, collecting EC-samples from manure storage and concentration sites may provide the optimum chance to detect MAP.

Repeated herd testing to determine MAP herd-status is frequently used in control programs, such as the Voluntary Bovine Johne's Disease Control Program (USDA, 2010) in the US. In this dataset, information from additional, consecutive EC-6 sets did not substantially change the estimated Se and Sp of EC. Numerically, the greatest increase in Se occurred with a second set of EC-6 samples, with minimal change in Se with additional sets of EC-6 (Table 2.5). Even though this was a relatively large study with 32 herds sampled repeatedly, when analyzing the data at the herd level, confidence intervals tended to still be very large, making it difficult to establish statistically significant differences between point estimates and predictors.

As the proportion of positive EC-samples within an EC-6 set increased, the probability increased for a cow within that herd to be PFC-IC positive (Figure 2.4). This cow-level probability is analogous to the WHP. When 1 of the 6 EC-samples in an EC-6 set was positive, estimated WHP was 0.7%, while when all 6 EC-samples were positive,

estimated WHP was 12.7%. The proportion of positive EC-samples within an EC-6 set provides an indication of MAP WHP. This knowledge is advantageous in MAP-control programs. For example, aggressive testing may be more beneficial and cost-effective in high-prevalence herds (Wells et al., 2003). When a herd had no positive EC-samples over two consecutive EC-6 samplings, the model predicted the probability of a cow being PFC-IC positive to be < 1%. Previous studies have estimated herd prevalence to be ≤ 2% when EC were negative (Pillars et al., 2009).

Based on estimated Se and Sp, environmental culture is an acceptable herd test for classification of MAP herd-status in MAP-negative and MAP-positive herds. Therefore, EC is an economical and non-invasive method of determining MAP herd-status. Knowledge of MAP herd-status and an estimate of WHP from EC can be valuable tools for herd MAP control programs.

2.6 References

Aly, S. S., R. J. Anderson, R. H. Whitlock, T. L. Fyock, S. McAdams, J. M. Adaska, J. Jiang, and I. A. Gardner. 2009. Reliability of environmental sampling to quantify Mycobacterium avium subspecies paratuberculosis on California free-stall dairies. J. Dairy Sci. 92:3634-3642.

Barkema, H. W., S. Hendrick, J. deBuck, S. Ghosh, G. G. Kaplan, and K. Rioux. 2011. Crohn's disease in humans and Johne's disease in cattle-Linked diseases? Pages 197-213 in Zoonotic Pathogens in the Food Chain. D. O. Krause and S. Hendrick ed. CAB International, Wallingford, UK.

Barkema H. W., J. W. Hesselink, S. L. B. McKenna, G. Benedictus, and H. Groenendaal. 2010. Global prevalence and economics of infection with Mycobacterium avium subsp. paratuberculosis in ruminants. Pages 10-21 in Paratuberculosis: organism, disease, control. M. A. Behr and D. M. Collins ed. CAB International, Wallingford, UK.

Berghaus, R. D., T. B. Farver, R. J. Anderson, C. C. Jaravata, and I. A. Gardner. 2006. Environmental sampling for detection of Mycobacterium avium ssp. paratuberculosis on large California dairies. J. Dairy Sci. 89:963-970.

Christensen, J. and I. A. Gardner. 2000. Herd-level interpretation of test results for epidemiologic studies of animal diseases. Prev. Vet. Med. 45:83-106.

Collins, M. T., I. A. Gardner, F. B. Garry, A. J. Roussel, and S. J. Wells. 2006. Consensus recommendations on diagnostic testing for the detection of paratuberculosis in cattle in the United States. J. Am. Vet. Med. Assoc. 229:1912-1919.

Collins, M. T., S. J. Wells, K. R. Petrini, J. E. Collins, R. D. Schultz, and R. H. Whitlock. 2005. Evaluation of five antibody detection tests for diagnosis of bovine paratuberculosis. Clin. Diagn. Lab. Immunol. 12:685-692.

Dohoo, I. R., S. W. Martin, and H. Stryhn. 2009. Veterinary Epidemiologic Research. 2nd ed. VER, Inc., Charlottetown, P.E.I., Canada.

Johnson-Ifearulundu, Y., J. B. Kaneene, and J. W. Lloyd. 1999. Herd-level economic analysis of the impact of paratuberculosis on dairy herds. J. Am. Vet. Med. Assoc. 214:822-825.

Kalis, C. H., M. T. Collins, H. W. Barkema, and J. W. Hesselink. 2004. Certification of herds as free of Mycobacterium paratuberculosis infection: Actual pooled faecal results versus certification model predictions. Prev. Vet. Med. 65:189-204.

Kim, S. G., E. H. Kim, C. J. Lafferty, L. J. Miller, H. J. Koo, S. M. Stehman, and S. J. Shin. 2004. Use of conventional and real-time polymerase chain reaction for confirmation of Mycobacterium avium subsp. paratuberculosis in a broth-based culture system ESP II. J. Vet. Diagn. Invest. 16:448-453.

Lavers, C., H.W. Barkema, I.R. Dohoo, S.L.B. McKenna, and G.P. Keefe. 2013. Evaluation of milk ELISAs for detection of Mycobacterium avium subspecies paratuberculosis in dairy herds and association with within-herd prevalence. J. Dairy Sci. In press.

Lombard, J. E. 2011. Epidemiology and economics of paratuberculosis. Vet. Clin. North Am. Food Anim. Pract. 27:525-535.

Lombard, J. E., B. A. Wagner, R. L. Smith, B. J. McCluskey, B. N. Harris, J. B. Payeur, F. B. Garry, and M. D. Salman. 2006. Evaluation of environmental sampling and culture to determine Mycobacterium avium subspecies paratuberculosis distribution and herd infection status on US dairy operations. J. Dairy Sci. 89:4163-4171.

Lombard, J. E., F. B. Garry, B. J. McCluskey, and B. A. Wagner. 2005. Risk of removal and effects on milk production associated with paratuberculosis status in dairy cows. J. Am. Vet. Med. Assoc. 227:1975-1981.

McKenna, S. L., G. P. Keefe, H. W. Barkema, and D. C. Sockett. 2005. Evaluation of three ELISAs for Mycobacterium avium subsp. paratuberculosis using tissue and fecal culture as comparison standards. Vet. Microbiol. 110:105-111.

Nielsen, S. S., M. A. Krogh, and C. Enevoldsen. 2009. Time to the occurrence of a decline in milk production in cows with various paratuberculosis antibody profiles. J. Dairy Sci. 92:149-155.

Pillars, R. B., D. L. Grooms, and J. B. Kaneene. 2009. Longitudinal study of the distribution of Mycobacterium avium subsp. paratuberculosis in the environment of dairy herds in the Michigan Johne's disease control demonstration herd project. Can. Vet. J. 50:1039-1046.

Pradhan, A. K. and Y. H. Schukken. 2009. Dynamics of endemic infectious diseases of animal and human importance on three dairy herds in the northeastern United States. J. Dairy Sci. 92:1811-1825.

Raizman, E. A., S. J. Wells, S. M. Godden, R. F. Bey, M. J. Oakes, D. C. Bentley, and K. E. Olsen. 2004. The distribution of Mycobacterium avium ssp. paratuberculosis in the environment surrounding Minnesota dairy farms. J. Dairy Sci. 87:2959-2966.

Smith, R. L., Y. H. Schukken, A. K. Pradhan, J. M. Smith, R. H. Whitlock, Van Kessel J.S., D. R. Wolfgang, and Y. T. Grohn. 2011. Environmental contamination with

Mycobacterium avium subsp. paratuberculosis in endemically infected dairy herds. Prev. Vet. Med. 102:1-9.

Tavornpanich, S., S. J. Wells, C. P. Fossler, A. J. Roussel, and I. A. Gardner. 2012. Evaluation of an alternative method of herd classification for infection with paratuberculosis in cattle herds in the United States. Am. J. Vet. Res. 73:248-256.

Tavornpanich, S., I. A. Gardner, R. J. Anderson, S. Shin, R. H. Whitlock, T. Fyock, J. M. Adaska, R. L. Walker, and S. K. Hietala. 2004. Evaluation of microbial culture of pooled fecal samples for detection of Mycobacterium avium subsp paratuberculosis in large dairy herds. Am. J. Vet. Res. 65:1061-1070.

Tiwari, A., J. A. VanLeeuwen, S. L. McKenna, G. P. Keefe, and H. W. Barkema. 2006. Johne's disease in Canada part I: Clinical symptoms, pathophysiology, diagnosis, and prevalence in dairy herds. Can. Vet. J. 47:874-882.

USDA. 2010. Uniform Program Standards for the Voluntary Bovine Johne's Disease Control Program. USDA-APHIS-VS, National Animal Health Monitoring System, Fort Collins, CO (APHIS 91-45-016).

Wells, S. J., S. M. Godden, C. J. Lindeman, and J. E. Collins. 2003. Evaluation of bacteriologic culture of individual and pooled fecal samples for detection of Mycobacterium paratuberculosis in dairy cattle herds. J. Am. Vet. Med. Assoc. 223:1022-1025.

Table 2.1. Definitions and acronyms describing terminology used in an evaluation of environmental fecal culture for the purpose of Mycobacterium avium subsp. paratuberculosis detection in dairy herds.

Term	Definition
Pooled fecal culture (PFC)	A pooled fecal sample consisting of individual fecal samples from 5 cows. Pools were created based on cows of similar age.
MAP-positive	A herd was considered MAP-positive if there was ≥ 1 positive PFC collected from the herd during the 2 yr project.
MAP-negative	A herd was considered MAP-negative if there were 0 positive PFC collected from the herd during the 2 yr project.
Individual cow fecal culture following a MAP positive-PFC (PFC-IC)	If a pooled fecal culture was positive, the frozen samples from the individual cows comprising that pooled fecal culture were thawed and cultured individually.
Apparent within-herd prevalence (WHP)	The apparent MAP within-herd prevalence was calculated based on the number of positive PFC-IC, divided by the number of cows from which fecal samples were collected in the herd.
Low-prevalence herd	A herd was considered to be MAP low-prevalence if there was ≤ 5% PFC-IC in a herd.
High-prevalence herd	A herd was considered to be MAP high-prevalence if there was > 5% PFC-IC in a herd.
EC-sample	One 95 ml jar of a composite environmental manure sample, collected from a specific location within the herd environment.
EC-6 (set)	A complete set of 6 EC-samples, collected from 6 specified areas within the herd environment.

Table 2.2. Summary of pooled fecal culture and environmental culture results for 15 herds with a minimum of one Mycobacterium avium subsp. paratuberculosis-positive fecal pool or environmental culture set.

Herd	Milking Herd Size	Pooled Fecal Culture (PFC)		Environmental Culture (EC)	
		Percent of herd visits with ≥1 positive PFC[b]	Mean WHP [a] (%)	Percent of herd visits with a positive EC-6 result[c]	Total positive EC-samples (%)
1	75	100 (3/3)	15.6	100 (6/6)	92
2	95	100 (3/3)	12.3	100 (6/6)	94
3	90	100 (3/3)	10.9	100 (6/6)	92
4	220	100 (3/3)	10.0	100 (6/6)	94
5	100	100 (3/3)	10.0	100 (6/6)	100
6	70	100 (4/4)	9.0	100 (6/6)	81
7	150	100 (3/3)	5.5	100 (6/6)	81
8	30	100 (4/4)	7.1	87 (7/8)	46
9	50	100 (4/4)	2.9	86 (6/7)	33
10	47	75 (3/4)	1.6	50 (4/8)	10
11	145	75 (3/4)	0.7	50 (4/8)	14
12	60	50 (2/4)	0.8	12 (1/8)	4
13	50	33 (1/3)	0.7	17 (1/6)	3
14	120	25 (1/4)	0.2	0 (0/8)	0
15	45	0 (0/4)	0.0	12 (1/8)	2

[a]Apparent within-herd prevalence (WHP) is based on PFC with culture of individual cow samples from positive pools.
[b]% of herd visits with ≥ 1 positive PFC (No. herd visits with ≥ 1 positive PFC/Total no. herd visits).
[c]% of herd visits with a positive EC-6 result, where EC-6 represents a set of 6 EC-samples collected from specified sites within the herd environment (No. positive EC-6 sets/Total no. EC-6 collections).

Table 2.3. Summary statistics for environmental culture samples collected and percent Mycobacterium avium subsp. paratuberculosis culture-positive by location on 32 dairy herds from the three Maritime provinces of Canada.

Location	Number of Samples	Percent Positive
Manure Storage (E.g. Lagoons, Manure Piles)	410	21.6%
Manure Concentration (E.g. Alleyways, Manure Gutters)	878	20.7%
Cow Concentration (E.g. Calving/Sick Pens)	98	8.2%
Total	1,386	20.4%

Table 2.4. Three logistic generalized estimating equation models for Mycobacterium avium subsp. paratuberculosis environmental culture sensitivity and specificity. Environmental culture result was the outcome in all models, with null models estimating overall sensitivity (MAP-positive herds) and specificity (MAP-negative herds) of environmental culture, and the multivariable model predicting the impact of fecal culture within-herd prevalence on the sensitivity of environmental culture.

Model	Estimate	95% Confidence Interval		P-value
		Lower	Upper	
Null logistic model of MAP-negative herds (Sp)				
(Intercept)	-4.93	-6.85	-3.02	
Null logistic model of MAP-positive herds (Se)				
(Intercept)	0.91	-0.03	1.86	
Multivariable logistic model of MAP-positive herds (Se)				
(Intercept)	-1.53	-2.40	-0.65	
Fecal culture test prevalence (proportion)	78.08	41.42	114.73	0.000

Table 2.5. Results of logistic generalized estimating equation models to evaluate sensitivity and specificity of environmental culture (EC), with increasing numbers of EC sets used to determine herd status.

	Models using MAP-positive Herds			Models using MAP-negative Herds		
No. of combined EC sets[a]	Intercept	#obs (#herds)	Sensitivity (95% CI)	Intercept	#obs (#herds)	Specificity (95% CI)
1	0.914	97 (14)	71% (49%-86%)	-4.93	140 (18)	99% (95%-100%)
2	1.45	83 (14)	81% (56%-93%)	-4.11	122 (18)	98% (90%-100%)
3	1.56	69 (14)	83% (57%-94%)	-3.95	104 (18)	98% (88%-100%)
4	1.64	55 (14)	84% (58%-95%)	-3.76	86 (18)	98% (86%-100%)
5	1.77	41 (14)	85% (59%-96%)	-3.53	68 (18)	97% (82%-100%)
6	2.24	27 (14)	90% (62%-98%)	-3.23	50 (18)	96% (72%-100%)

[a]An EC set consisted of 6 samples collected from 6 sites within the farm environment.

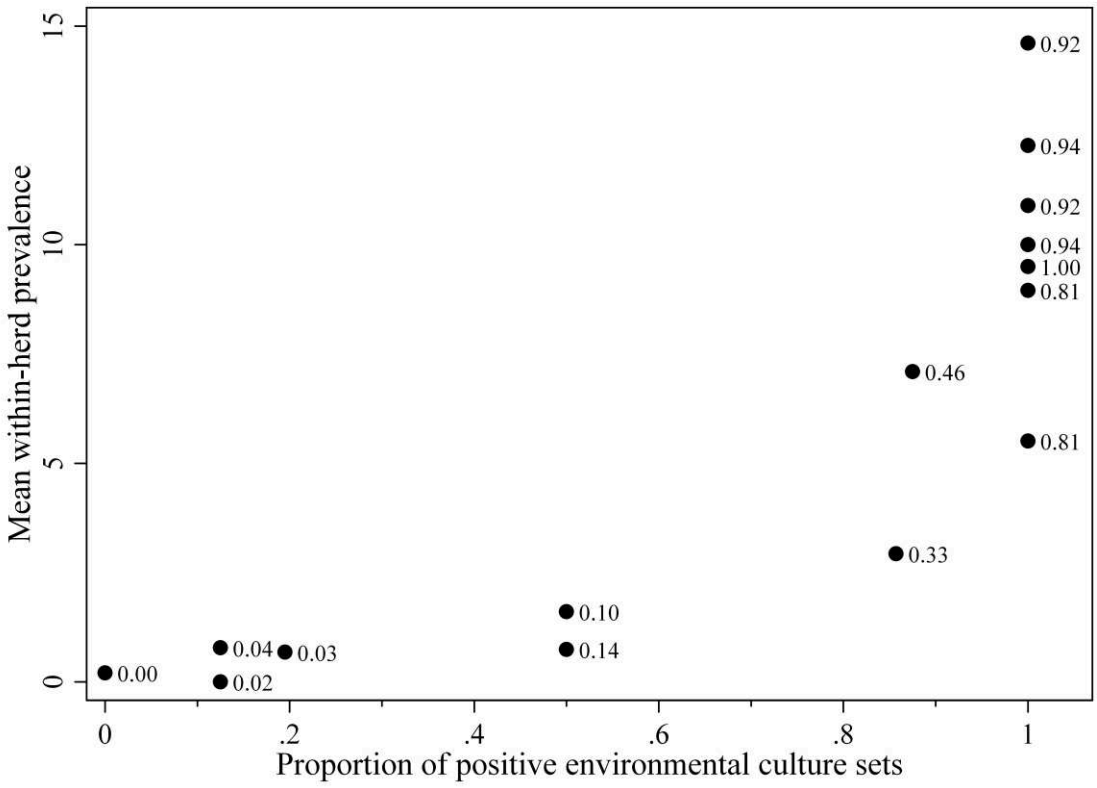

Figure 2.1. Mean apparent within-herd Mycobacterium avium subsp. paratuberculosis fecal culture prevalence[a] compared to the proportion of positive environmental culture sets in 15 dairy herds that were fecal culture and/or environmental culture-positive. Number markers beside data points indicate the proportion of environmental culture samples that were positive within all environmental culture sets collected from that herd.

[a]Apparent within-herd prevalence is based on pooled fecal culture with culture of individual cow samples from positive pools.

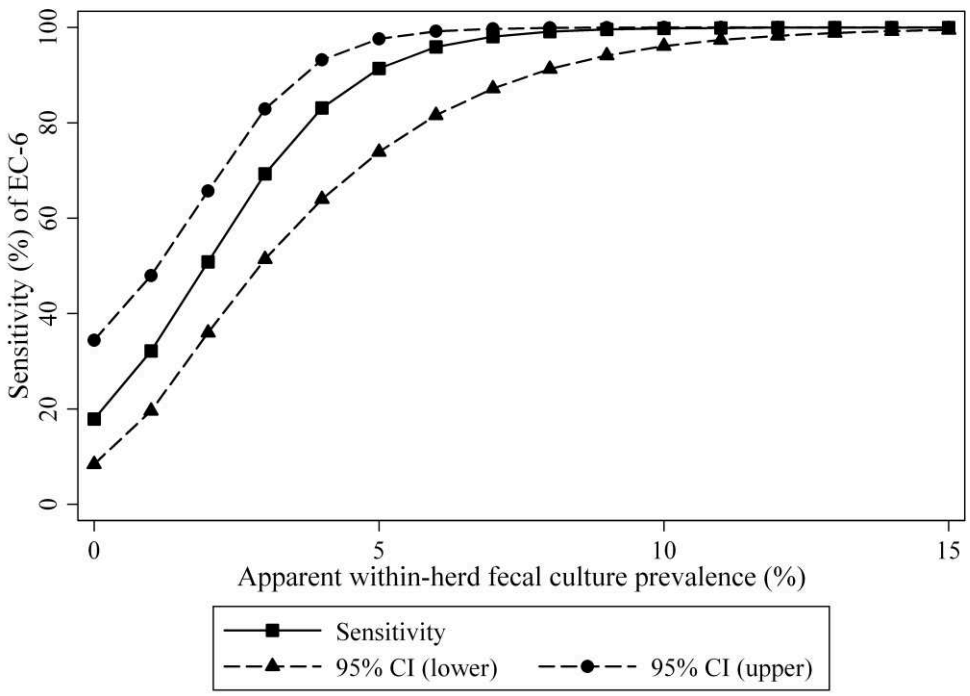

Figure 2.2. Sensitivity and 95% confidence intervals (CI) of single sets of 6 environmental culture samples (EC-6) by apparent Mycobacterium avium subsp. paratuberculosis within-herd test prevalence[a], using a logistic generalized estimating equations model.

[a]Apparent within-herd prevalence is based on pooled fecal culture with culture of individual cow samples from positive pools.

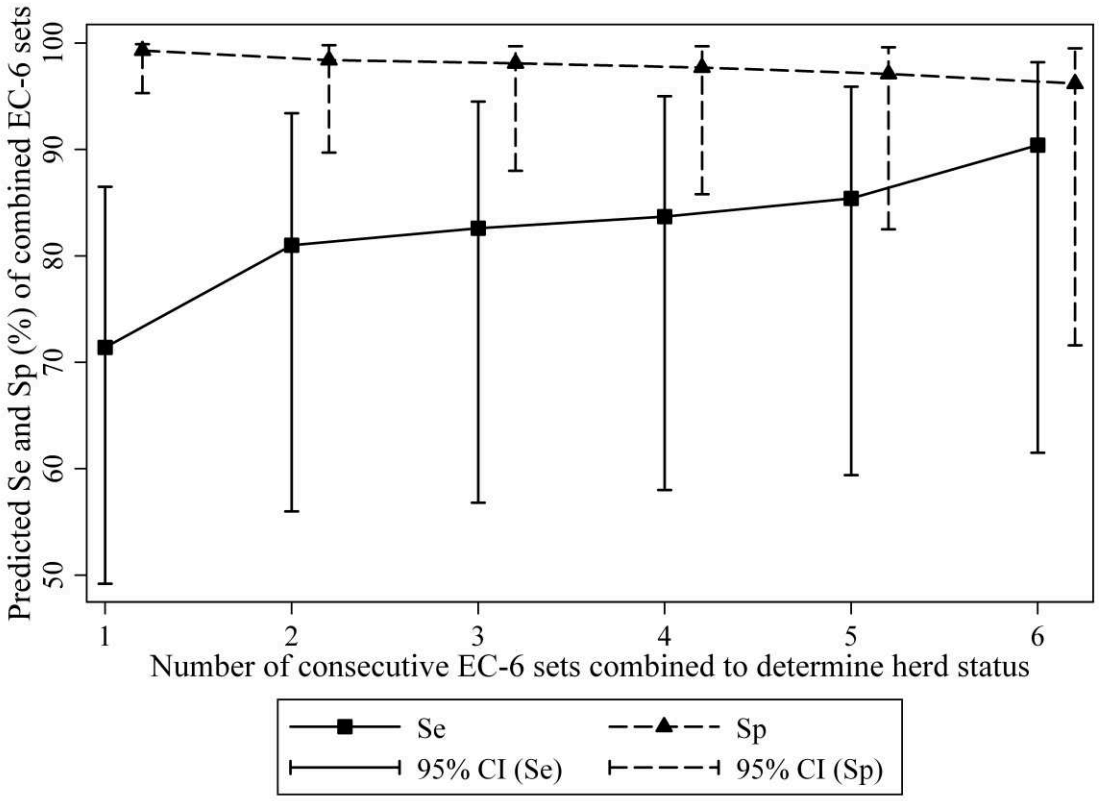

Figure 2.3. Predicted sensitivity (Se) and specificity (Sp), with 95% confidence intervals (CI), of sets of 6 environmental culture samples (EC-6) for Mycobacterium avium subsp. paratuberculosis, by the number of EC-6 sets used to determine herd status.

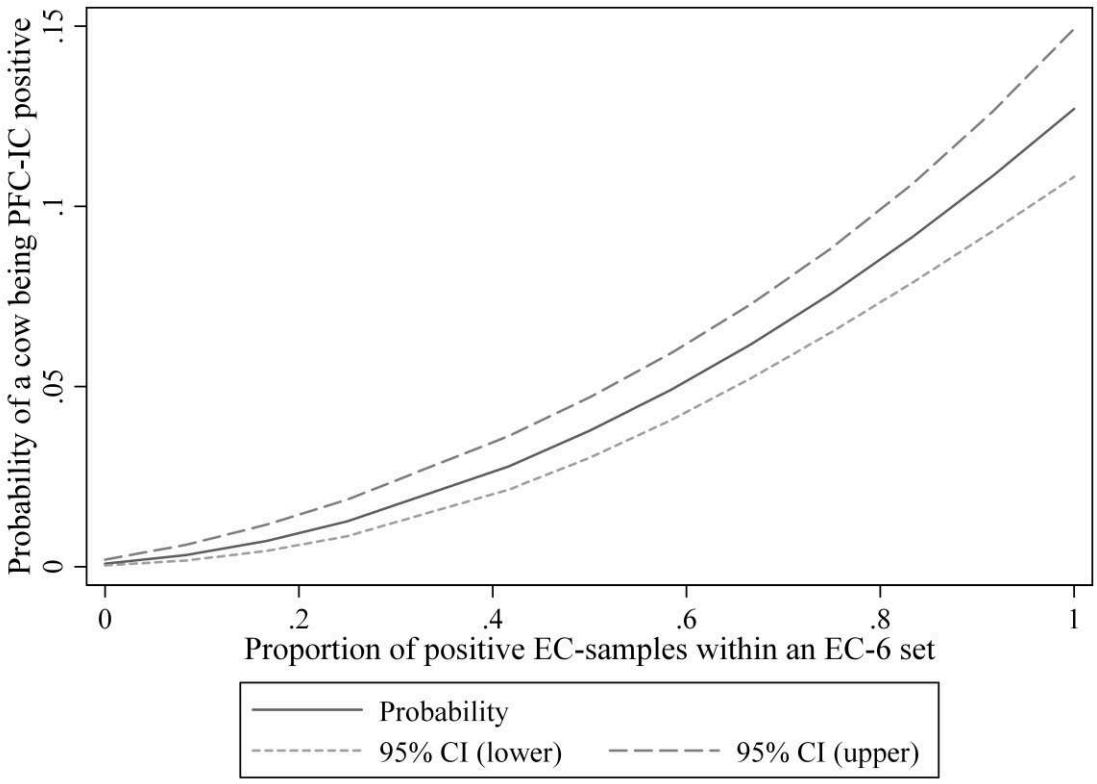

Figure 2.4. The effect of the proportion of positive Mycobacterium avium subsp. paratuberculosis environmental culture samples on the probability[a] of a cow within a herd being positive based on pooled fecal culture with individual fecal culture follow-up (PFC-IC).

[a]The probability of a cow to be PFC-IC positive is equal for each cow within the herd, and is analogous to the apparent within-herd prevalence.

CHAPTER 3

EVALUATION OF MILK ELISAS FOR DETECTION OF MYCOBACTERIUM AVIUM SUBSPECIES PARATUBERCULOSIS IN DAIRY HERDS AND ASSOCIATION WITH WITHIN-HERD PREVALENCE

Carrie J. Lavers, Shawn L. B. McKenna, Ian R. Dohoo, Herman W. Barkema, Greg P. Keefe

The majority of contents of this chapter are published as: C. J. Lavers, H. W. Barkema, I. R. Dohoo, S. L. B. McKenna, and G. P. Keefe. 2013. Evaluation of milk ELISAs for detection of Mycobacterium avium subspecies paratuberculosis in dairy herds and association with within-herd prevalence. J. Dairy Sci. In press.

3.1 Abstract

Cow-level milk ELISA results can be used to determine herd Mycobacterium avium subsp. paratuberculosis (MAP) status. Milk sample collection is minimally invasive, and ELISA results can be obtained quickly and economically. The study objectives were to evaluate the test characteristics of 3 commercial milk ELISAs conducted on individual cows but interpreted at the herd level, and to determine the impact of within-herd MAP prevalence on the performance of the milk ELISA herd test. A total of 32 purposively selected herds with a median herd size of 66 milking cows were used in this 2 yr project. Fecal and milk samples were collected from all milking cows at 6 mo intervals. Fecal samples were pooled by cow-age, with 5 cow samples per pool. Subsequent individual fecal culture was completed on cow samples from positive pools. Herd MAP-status was defined as MAP-positive if, at any point during the longitudinal study, a pooled fecal culture from the herd was positive. Milk samples were analyzed using each of 3 commercial milk ELISA kits; a cow-level result from each ELISA was classified as positive following the respective manufacturer's recommended threshold for a positive result. Herd-level milk ELISA test characteristics were estimated using GEE logistic models, which accounted for repeated measurements. Using a cutoff of 2% milk ELISA-positive cows, milk ELISA herd sensitivities were: ELISA A: 59% (95% CI: 36-78%); ELISA B: 56% (95% CI: 32-77%); and ELISA C: 63% (95% CI: 41-81%). Herd specificities for ELISA A, B, and C were 80% (95% CI: 71-88%), 96% (95% CI: 89-98%) and 92% (95% CI: 86-96%), respectively. The remainder of the statistical analyses focused on results from ELISA B. Although herd sensitivity of ELISA B increased as

within-herd prevalence increased, in low prevalence herds the herd sensitivity of the milk ELISA was low. As examples, herd sensitivity was 11% when within-herd prevalence was 1%, and 62% when within-herd prevalence was 5%. Categorical likelihood ratios based on positive milk ELISA test results predicted that herds with a milk ELISA within-herd prevalence above 0 but < 2% had a similar likelihood to be MAP-positive or MAP-negative, whereas herds with a milk ELISA prevalence between 2 and 4% were 3.7 times more likely to be MAP-positive than MAP-negative. All herds with a milk ELISA prevalence > 4% were MAP-positive. Although milk ELISA B worked well to establish herd MAP-status in high-prevalence herds, interpretation was unreliable in MAP-negative and low-prevalence herds.

3.2 Introduction

Mycobacterium avium subspecies paratuberculosis (**MAP**) is the causative agent of Johne's disease (**JD**). Although eradication of MAP is very difficult, best management practices that limit calf exposure to MAP combined with testing programs that result in management changes for, or culling of, positive cows, have decreased within-herd prevalence over time (Collins et al., 2010). While knowledge of herd MAP-status is not required for best management practices to be implemented, determining that a herd is MAP-positive provides further motivation for producers to adopt management changes to decrease both within- and between-herd transmission of the bacterium.

Testing options for determination of herd MAP-status include polymerase chain reaction (**PCR**) or culture on individual or pooled cow fecal samples or pooled

environmental samples, and ELISA on milk or serum. Performing individual cow fecal cultures on an entire herd is generally considered cost-prohibitive for determining herd MAP-status, as fecal culture is the most expensive cow-level test (Tiwari et al., 2006). Individual fecal culture specificity approaches 100% (Nielsen and Toft, 2008), but sensitivity is limited, with estimates ranging from 23% (McKenna et al., 2005) to 29% (Whitlock et al., 2000) for infected and 74 to 89% for infectious animals, depending on the culture method used (Sockett et al., 1992). Pooling individual cow fecal samples by cow-age, with 5 cows per pooled fecal culture (**PFC**), substantially decreases cost and is reported to be both highly sensitive and specific, relative to individual culture, for determination of herd MAP-status (Kalis et al., 2004). Wells et al. (2003) reported that PFC detected 16 of the 17 herds detected with individual fecal culture, equating to a herd-level sensitivity (**HSe**), relative to individual culture, of 94%. Nine of the 17 herds had a prevalence of infection, based on individual fecal culture, < 10%, and the 1 herd detected positive by individual culture, but negative with PFC, had only a single light shedding cow. In the same study, herd-level specificity (**HSp**) of 100%, relative to individual fecal culture, was reported (Wells et al., 2003). Another method of determining herd MAP-status is to test individual lactating cows with a milk ELISA and compile results at the herd-level. Milk ELISA herd testing is both inexpensive and rapid compared to fecal culture, and is easy to implement in dairy herds which participate in DHI programs involving monthly individual cow milk collection.

The majority of MAP-positive herds have a low to moderate within-herd prevalence, and although published herd-prevalence estimates vary, MAP-negative herds are expected to represent a portion of dairy herds in most regions (Lombard, 2011).

Because these low-prevalence and MAP-negative herds represent a majority of the dairy industry, and HSe depends on within-herd prevalence (Dohoo et al., 2009), it is important to determine how the milk ELISA herd test performs in these herds. Previous studies have estimated an HSe for the milk ELISA, but did not take into account the within-herd prevalence levels of the herds tested (Hendrick et al., 2005; Lombard et al., 2006), also HSp estimates have been rare (Hendrick et al., 2005).

The objectives of this study were to evaluate the herd-level test characteristics of 3 commercial milk ELISAs, conducted on individual cows but interpreted at the herd level, and to determine the impact of within-herd MAP prevalence on the performance of the milk ELISA herd test.

3.3 Materials and Methods

3.3.1 Purpose for Testing and Study Terminology

The purpose for testing was to classify a dairy herd as MAP-infected or not infected (Collins et al., 2006). This establishment of whether MAP infection is present in a dairy herd or not is the initial step in a herd risk management plan for MAP control within a herd (Garry, 2011). The target condition was a MAP-infected herd, which is a herd that has at least 1 MAP-infected animal (an animal that has MAP in its tissues) (Gardner et al., 2011). For the purpose of this study, the target condition was not expected to change over the period of time herds were tested. A case definition provides a practical definition of the target condition using a reference standard (Gardner et al., 2011). Nielsen and Toft (2008) suggested that, although not perfect, the best available

method to establish cow infection status would be post-mortem histopathological evaluation of up to 100 tissues per cow. This would be extremely expensive (Nielsen and Toft, 2008), and cannot be accomplished on live cows, so a case definition was used of a MAP-infected herd being a herd where MAP is shed in feces, as identified by having 1 or more positive PFC. In order to strengthen PFC as a herd reference standard, repeated (3 to 4) herd fecal collections were used. The case definition is referred to as a MAP-positive herd, and this term, along with others defined in the manuscript, is denoted in Table 3.1. The case definition is similar to the use of PFC in the Uniform Program Standards for the Voluntary Bovine Johne's Disease Control Program in the USA (USDA, 2010), where a herd with ≥ 1 positive pooled fecal sample is considered an infected herd. The reference standard test chosen for this project was repeated PFC. It is recognized that there is no true and perfect 'gold standard' diagnostic test for MAP at either the cow or herd-level (Gardner et al., 2011). With a HSe approaching that of individual fecal culture (Wells et al., 2003; Kalis et al., 2004; Collins et al., 2006), PFC was repeated over the project to minimize the misclassification of low-prevalence herds as MAP-negative, which is recognized as an issue when attempting to detect MAP-infected herds (Wells et al., 2002; Kalis et al., 2004).

3.3.2 Study Design

A total of 34 Canadian herds from three Canadian provinces, Prince Edward Island, New Brunswick and Nova Scotia, participated in this prospective study from June, 2009 to January, 2011. Details on herd selection, herd MAP-status classification, and herd demographics have been reported (Lavers et al., 2013). In short, the goal of the purposive herd selection was to obtain a mixture of high-prevalence, low-prevalence, and

MAP-negative herds. Collection of milk and fecal samples from all lactating cows was done at 0, 6, 12, and 18 mo (4 rounds of testing). The initial round of herd testing was completed on 27 herds. Because there were too few MAP-positive herds, 7 herds were added for the remaining 3 rounds of herd testing. These herds were chosen based on the occurrence of 1 or more cows with either a positive individual fecal culture (n=4) or clinical signs suggestive of paratuberculosis and a positive serum or milk ELISA test (n=3) in the previous 2 yr. Two of the 34 herds were subsequently excluded for compliance reasons because they did not have at least 3 herd visits with collection of individual cow fecal samples. As a result, of the 32 herds included in the analyses, 25 had 4 rounds of testing and 7 herds had 3 rounds of testing. Median herd size was 66 milking cows (mean: 82; range: 28 to 220). The herd median cow age at testing ranged from 2.9 to 5.5 y for the 32 herds. Considering all cows tested from all herds, the median age at testing was 4.0 y (mean: 4.4 y; range: 1.8 – 17.3 y). One herd expanded from 190 to 220 lactating cows during the project. This herd had a PFC-positive test prior to the introduction of new animals. The remaining study herds were not necessarily closed, but they did not have substantial introductions of new animals during the project.

3.3.3 Sample Collection

Fecal sample collection has been described in Lavers et al. (2013). Briefly, individual fecal samples were collected at 6 mo intervals from all lactating cows, using a clean full-length plastic glove lubricated with sterile water. If samples could not be processed immediately, they were frozen at -20 or -80˚C if processing would occur within either 2 wk or between 2 and 6 wk after collection, respectively. Bronopol-preserved individual milk samples were collected at 6 mo intervals from all lactating

cows as part of routine herd milk testing by the regional DHI organization (Valacta, Montreal, QC, Canada). The aim was to collect fecal and milk samples from each round as close as possible to each other in time, with no more than 3 mo between a herd's milk and fecal sample collection.

3.3.4 Laboratory Testing: Fecal Culture

The technical laboratory staff at the Maritime Quality Milk Laboratory (Atlantic Veterinary College, Charlottetown, PEI, Canada) processed all fecal samples. This laboratory was approved by USDA proficiency-testing for this technique. Pooled fecal samples were created by pooling individual samples by cow age, with 5 cows in each PFC (Kalis et al., 2000). To estimate cow MAP-status, individual cow samples from positive-PFC were thawed and cultured individually. Fecal cultures were performed using ESP para-JEM® broth (Nova Century Scientific, Inc., Burlington, ON, Canada) according to manufacturer's protocol, with the exception that samples were incubated for 49 d, rather than 42 d. Briefly, 2 g of feces from each individual cow sample was put into a weigh boat and the resulting 10 g of pooled feces was mixed thoroughly with a wooden tongue depressor, then 2 g of the mixed sample was added to 35 mL of sterile water in a 50 mL centrifuge tube. For individual cow samples, 2 g of manure was measured and placed into 35 mL of sterile water in a 50 mL centrifuge tube. Samples were agitated vigorously for 10 min, and then placed upright for 30 min. A 5 mL sample was drawn from the feces and water mixture and placed into a centrifuge tube containing 25 mL of 0.9% 1-hexadecylpyridinium chloride in half-strength brain heart infusion broth. The tubes were incubated for 18-24 h at 35-37^{o}C, followed by centrifugation at 3000 rpm for 20 min. The supernatant was discarded and the pellet was re-suspended in 1 mL of brain

heart infusion broth and antibiotic mixture (50 µg amphotericin B, 100 µg vancomycin and 100 µg naladixic acid). Tubes were incubated for 24 h at 35-37°C. On Day 3, 2.5 mL of a pre-mixed para-JEM® reagent (1 mL egg yolk supplement, 1 mL growth supplement and 0.5 mL antibiotic supplement) and 1 mL of the decontaminated sample were added to a para-JEM® broth bottle. The broth bottle was agitated and then placed in the ESP® Culture System II (TREK Diagnostic Systems, Inc., Cleveland, OH, USA) for incubation up to a maximum of 49 d. The culture system's computer output was examined daily for indications of positive curves.

Confirmatory tests were performed as per the protocol outlined in Lavers et al. (2013). Briefly, when headspace pressure indicted growth or at the end of the 49 d incubation period, acid-fast staining was conducted on the culture broth. Presumptive positive samples (acid-fast positive or positive growth curves) were selected for confirmation with the VetAlert™ Johne's Real-Time PCR kit (Tetracore, Inc., Rockville, MD, USA), which targeted the hspX gene. A culture sample was considered positive if the presumptive positive sample was positive based on PCR testing.

3.3.5 Laboratory Testing: Milk ELISA

Three indirect ELISAs were used: Parachek2 Mycobacterium paratuberculosis test kit® (Prionics AG, Schlieren-Zurich, Switzerland; **ELISA A**), Mycobacterium paratuberculosis antibody test kit® (IDEXX, Westbrook, ME, USA; **ELISA B**), and Paratuberculosis Indirect® (IDVet Innovative Diagnostics, Montpellier, France; **ELISA C**). All 3 commercial milk ELISAs were performed by the technical laboratory staff at the Maritime Quality Milk laboratory (Charlottetown, PEI, Canada), which was USDA proficiency-tested for ELISA A. All 3 kits used an absorption step to remove antibodies

that cross-react with Mycobacterium phlei (Yokomizo et al., 1985). Milk ELISAs were performed following manufacturers' instructions; a cow milk sample was defined as ELISA-positive based on the criteria provided in the specific kit.

3.3.6 Statistical Analyses

Descriptive statistics were calculated for proportions of positive test results, in total and among MAP-positive and MAP-negative herds. Unless otherwise specified, for the test under evaluation, a milk ELISA-positive herd was a herd with ≥ 2% of cows milk ELISA-positive within the herd. This milk ELISA cutoff was previously used by Lombard et al. (2006), and receiver operating characteristic evaluations indicated a balance between HSe and HSp with this cutoff value.

Herd sensitivity and HSp were determined, using generalized estimating equations (**GEE**) with an exchangeable correlation structure, to account for the repeated nature of the data. Milk ELISA HSe was estimated from a null logistic GEE model with milk ELISA herd test results from MAP-positive herds as the outcome, using the equation: $HSe=e^{\beta o}/(1+e^{\beta o})$, where β_0 was the intercept coefficient of a null logistic GEE model based on MAP-positive herds only (Dohoo et al., 2009). Similarly, milk ELISA HSp was estimated from the intercept coefficient of a null logistic GEE model with milk ELISA herd test results on MAP-negative herds, using the equation $HSp=1-(e^{\beta o}/(1+e^{\beta o}))$. In addition to the cutoff of 2% milk ELISA-positive cows, HSe and HSp were also calculated using cutoffs of 1 and 3% milk ELISA-positive cows.

For evaluation of factors influencing HSe and HSp, estimation of predictive values and determination of likelihood ratios (**LR**), results for ELISA B only are

presented. This ELISA had a numerically superior HSp compared to ELISA A, and is more commonly used in North America than ELISA C.

In order to evaluate predictors influencing HSe and HSp of milk ELISA B, logistic GEE models were created. The initial model for HSe, for example, was:

$$\text{logit}(p) = \beta_0 + \text{SEASON} + \text{HSIZE} + \text{AGE} + \text{WHP}$$

where p is the probability of a MAP-positive herd to have a positive milk ELISA herd test result; β_0 is the common intercept; SEASON is a categorical variable indicating season of milk sample collection; HSIZE is the herd size (continuous); AGE is the median herd cow-age (continuous); and WHP is the mean FC within-herd MAP prevalence (continuous). As defined in Table 3.1, WHP was calculated as the number of fecal culture-positive cows divided by the total number of cows from which fecal samples were collected. A herd's mean WHP was the mean of the herd's WHP estimates calculated from each herd collection. Predictors with a P-value ≤ 0.15 in univariable analyses were entered into a multivariable logistic regression GEE model. Furthermore, P ≤ 0.05 was considered significant for inclusion in the final model. Lowess smoothers and fractional polynomials were produced to investigate linearity and to explore power transformations of significant predictors in an effort to optimize linearity, respectively. Based on these investigations, no transformation of predictors was required.

The probability for a herd to be MAP-positive, given a milk ELISA positive (positive predictive value of a positive test) or milk ELISA-negative herd test (positive predictive value of a negative test) (Dohoo et al., 2009) was calculated for a range of

within and between herd prevalence scenarios for ELISA B at a 2% cutoff. The HSp estimate used in the calculation of these probabilities was generated in the null GEE model. The multivariable GEE model used to estimate HSe was influenced by the WHP; therefore, probabilities were calculated using HSe estimates from herds with WHP of 2 and 10%, to represent low and high WHP levels, respectively. Probabilities were calculated over a range of test population herd prevalence estimates.

Categorical LR were determined for 4 categories of milk ELISA B prevalence (0%, > 0 to 2%, > 2 to 4%, and > 4%). Within each milk ELISA prevalence category, the LR was calculated as the proportion of MAP-positive herds within the category divided by the proportion of MAP-negative herds within the category.

Stata/IC® Version 11.2 (StataCorp LP, College Station, TX, USA) was the statistical software used for all statistical analyses. A P-value ≤ 0.05 was considered significant.

3.4 Results

3.4.1 Herd Prevalence

Based on PFC results collected over the 2 yr study period, 14 (44%) herds were MAP-positive, whereas 18 (56%) herds were MAP-negative. Nine of the 14 MAP-positive herds had at least 1 PFC-positive result at each round of sampling, whereas 5 MAP-positive herds fluctuated between PFC test-negative results and at least 1 PFC-positive result (Table 3.2). In the first year of testing (Rounds 1 and 2), 5 of 9 herd tests in these 5 fluctuating MAP-positive herds had at least 1 cow fecal sample that was

culture-positive. Similarly, in the second year of testing (Rounds 3 and 4), 5 of 10 herd tests in these 5 fluctuating MAP-positive herds had at least 1 cow fecal sample that was culture-positive (Table 3.2).

In the 18 MAP-negative herds, 29, 13 and 18% of the 72 milk ELISA herd tests had at least 1 milk ELISA-positive cow for ELISA A, B and C, respectively (Table 3.3). In the 14 MAP-positive herds, 71, 67 and 78% of the 49 milk ELISA herd tests had at least 1 cow milk ELISA-positive for ELISA A, B and C, respectively (Table 3.3).

3.4.2 Within-Herd Prevalence

Mean within-herd WHP ranged from 0 to 15.6% (Fig. 3.1). Seven of the 18 MAP-negative herds had at least 1 milk ELISA B-positive cow sample. In the 14 MAP-positive herds, the WHP was generally higher than milk ELISA prevalence estimates (Fig. 3.1); mean fecal culture WHP was 6.2%, whereas mean milk ELISA prevalence was lower, being 2.9, 3.0, and 4.0% for ELISA A, B, and C, respectively.

3.4.3 Milk ELISA Test Characteristics

Based on a null logistic GEE model of MAP-positive herds, milk ELISA HSe ranged from 58 to 63% for the 3 milk ELISAs (Table 3.4). Based on a null logistic GEE model of MAP-negative herds, HSp ranged from 80 to 96%.

Results for modeling impact of predictors on HSe and HSp, and estimation of LR and predictive values, refer to ELISA B only. In the final multivariable logistic GEE model evaluating predictors influencing milk ELISA HSe, fecal culture WHP was a significant positive predictor (Fig. 3.2), and no other predictors were significant. No factors were found to be significant predictors of milk ELISA B HSp. The intercept

coefficient was -2.68, and the coefficient for WHP (%) was 0.63 (P = 0.001). As examples, milk ELISA HSe was 19% when WHP was 2%, and increased to 97% when WHP was 10%.

The probability of a herd being MAP-positive, given a milk ELISA positive herd test, increased as both WHP and herd prevalence in the test population increased (Fig. 3.3). For example, in a test population with a 20% herd prevalence, the probability of being MAP-positive, given a milk ELISA positive herd test was 54% in herds with a 19% HSe (the expected HSe when WHP within the herd is 2%). Conversely, in the same population of herds, the probability of being MAP-positive was 85% when the HSe was 97% (the expected HSe when WHP within the herd is 10%). In contrast, the probability of a herd being MAP-positive, given a milk ELISA negative herd test, decreased as HSe decreased. For example, in the same population of herds with a herd prevalence of 20%, at a HSe of 19% (the expected HSe when WHP within the herd is 2%), the probability of a herd being MAP-positive, given a milk ELISA negative herd test, was 17%. Conversely, in the same population of herds the probability of being MAP-positive was only 1% when the HSe was 97% (the expected HSe when WHP within the herd is 10%). Likelihood ratios for 4 categories of milk ELISA prevalence ranged from 0.3 to ∞ (Table 3.5).

3.5 Discussion

Within-herd prevalence can have a substantial effect on test characteristics at the herd level (Dohoo et al., 2009), and has apparently not been evaluated previously in a

milk ELISA study. The objective of this study was to both evaluate the herd-level test characteristics of milk ELISAs, as well as to determine the impact of within-herd MAP prevalence on the performance of the milk ELISA herd test. Herd sensitivity and HSp were estimated for 3 commercial milk ELISAs, and WHP had a significant impact on HSe.

In the present study of 44% MAP-positive herds, HSe of the 3 milk ELISAs ranged from 56 - 63%, using the 2% prevalence milk ELISA cutoff defining a positive herd. Klausen et al. (2003) predicted a milk ELISA HSe of 83%, based on a single positive cow cutoff and data collected from 6 high-prevalence and 2 negative herds. However, they suggested that sampling from a more representative set of herds was required for a less biased estimate of test characteristics. Comparison of current study results to Klausen et al. (2003) is limited because of the differing within-herd prevalence estimates between study herds of the 2 projects. Hendrick et al. (2005) estimated milk ELISA HSe ranged from 87 - 92%, depending on whether a 2 or 1 positive cow cutoff, respectively, was used. However, the Hendrick et al. (2005) estimates of HSe were higher than those in the current study because those HSe estimates were derived from epidemiologic formulae, rather than comparison to a reference standard. In contrast, Lombard et al. (2006) used a field application of the milk ELISA herd test compared to individual fecal culture, and reported HSe values for the milk ELISA of 61% for a 2% cutoff and 83% for a 1 positive cow cutoff. Those HSe estimates were similar to the current study. However, Lombard et al. (2006) did not evaluate the impact of WHP on milk ELISA HSe.

Based on modeling milk ELISA HSe over a range of WHP values, HSe increased as WHP increased. The contrast in HSe (Fig. 3.2), from less than 20% in herds with 2% WHP to more than 95% in herds with 10% WHP, indicates that this is a critical factor. Wells et al. (2002) considered the effect of within-herd prevalence in their serum ELISA HSe estimate, using a single positive cow cutoff, and also noted an increase in HSe from 91 to 100% across 3 categories of increasing within-herd fecal culture prevalence (< 5%; 5 - 10%; and ≥ 10%). Although the trend is the same, the HSe estimates were higher than those from the current study. In the Wells et al. (2002) study, 41% of the herds had a within-herd prevalence, based on individual fecal culture, of ≥ 10%, while in the current study 21% of herds had a WHP ≥ 10%. The overall lower mean HSe estimate in the current study as compared to Wells et al. (2002), as well as milk ELISA test characteristics estimated by Hendrick et al. (2005), could also be attributed to inclusion of low-prevalence herds in the current study population. Lombard et al. (2006) estimated a similar overall HSe to the current study, reflecting a similarly greater proportion of study herds that were low-prevalence. Previous publications have reported that a solitary, low-shedding cow in a pool may not be detected (Wells et al., 2003). If this occurred in the current study, WHP estimates may have been an underestimation of true within-herd MAP prevalence. Based on the low HSe in low-prevalence herds estimated from the current study, we inferred that whereas the milk ELISA herd test may be useful for a control program in a high-prevalence herd, as suggested by consensus recommendations (Collins et al., 2006), HSe is too low to detect a low-prevalence herd.

Herd sensitivity will increase as the cutoff is lowered, but this increase will come at a cost to the herd specificity (Martin et al., 1992). While previous studies have used a 1

cow cutoff (Hendrick et al., 2005; Lombard et al., 2006), the lower cutoff translates to a lower HSp. In the current data, between 13 and 29% of the milk ELISA herd tests from MAP-negative herds had at least 1 positive ELISA result. For the purposes of applying the milk ELISA test in MAP control programs, this number of false-positive herd results was deemed to be too high. In the present study, a 2% positive cow cutoff was used because it allowed for a more appropriate balance between HSe and HSp, as well as for a greater extension to a variety of herd sizes.

In light of the potential impact of herd characteristics (within-herd prevalence, size) on validity of test characteristic estimates, representativeness of study herds has been previously called into question (Klausen et al., 2003). To address the issue of variable HSe, herds used in this study were chosen to represent a range of within-herd prevalence values (range from 0 to 15.6%). Additionally, herd size is another important aspect of representativeness. Study herds were similar in size to Canadian (CDIC, 2013) and European Union (DairyCo, 2012) norms. Herd size in the United States is highly variable; whereas 68% of cows are housed on farms with > 200 cows, 82% of the dairy herds with ≥ 30 cows in the United States have less than 200 cows (USDA, 2012). The variability in HSe (Fig. 3.2) indicates that concern over test characteristic variability across herds was warranted, and our study design allowed for generalization of results across a large proportion of herds in the North American and European dairy industries.

Previous studies have used herds with a history of clinical disease or high-prevalence of disease (Wells et al., 2002; Klausen et al., 2003). In this study, we attempted to select herds with a wider range of within-herd prevalence, although they were not randomly selected. This was done in order that the study population would more

closely reflect within-herd prevalence expected in the dairy industry (target population) (Lombard, 2011). However, additional research using a larger, randomly selected, sample of herds would be useful.

In the present study, HSp of the 3 milk ELISAs ranged from 80 - 96%. In other studies, HSp could not reliably be estimated, because either few or no MAP-negative herds were included, or the herd MAP-status was unknown. Klausen et al. (2003) estimated an HSp of 100%, but had only 2 negative herds. Using calculated estimates, without a known herd MAP-status based on fecal culture, Hendrick et al. (2005) estimated an HSp of 66 and 93% for a cutoff of 1 or 2 positive cows, respectively. Lombard et al. (2006) had only 3 fecal culture test-negative herds with milk ELISA data and did not estimate an HSp. Wells et al. (2002) included 8 herds that had no prior clinical or laboratory evidence of paratuberculosis; although they did not formally calculate an HSp, all 8 herds were classified as MAP-positive with the serum ELISA. These studies did not have a large number of herds that were MAP-negative at several consecutive tests from which to calculate HSp. With 18 herds that were repeatedly PFC-negative, the present study provided a strong reference group from which to estimate HSp.

Compared to a single cross-sectional herd PFC, repeated testing with PFC allowed for the case definition to more closely reflect the target condition of a MAP-infected herd, because repeated sampling increases the probability to identify low prevalence infected herds (Wells et al., 2002; Kalis et al., 2004), and is recommended for establishment of a MAP-negative herd status in control programs (Collins et al., 2005). In this study, test characteristics were estimated from a well-characterized group of 32 herds

over 2 yr. Nine of these herds were consistently positive on PFC (100% of the 30 PFC herd tests collected from these 9 herds were positive), 5 were intermittently positive on PFC (low-prevalence) and 18 herds were repeatedly PFC negative and were considered MAP-negative. These low-prevalence herds were positive on PFC for 53% of the herd PFC tests (From Table 3.2, 10/19 herd PFC tests were positive). If only a single cross-sectional sampling was used as a case definition, these low-prevalence herds that were intermittently positive on PFC could have been misclassified as a MAP-negative herd, potentially resulting in a lower reported milk ELISA HSp if misclassified herds also had positive milk ELISA herd tests.

Regarding misclassification of a MAP non-infected herd, fecal culture specificity is expected to approach 100%, so a false-positive PFC, resulting in a false-positive herd diagnosis, is unlikely. There were 2 herds, classified MAP-positive, that had a single positive PFC throughout the study. Although improbable, it is possible that the shedding cow was only transiently in the herd, and that no other cows in the herd were MAP-infected. Given the low expected probability for this scenario, potential impact on test evaluation outcomes is expected to be minimal. The milk ELISA herd test could be evaluated against other diagnostic herd test outcomes to ensure similar results. An alternative to the current PFC-based case definition model would be individual culture of fecal samples. Although PFC has reported high HSe relative to individual cow fecal culture (Wells et al., 2003), repeated individual cow fecal cultures may have allowed for an improved case definition, but was cost-prohibitive. A Bayesian analysis, which allows for analysis of test characteristics without requiring a comparative ‘gold standard’, would also have circumvented the challenge of an imperfect ‘gold standard’ herd test. However,

a sample size beyond the financial scope of this project would be required to elicit reasonable estimates and credibility intervals from a Bayesian analysis.

Both herd prevalence and herd-level test characteristics can profoundly influence interpretation of milk ELISA herd test results in low-prevalence herds (Fig. 3.3). For practical field applications, it is important to understand the probability that a herd does or does not actually have MAP based on test results. At the cow-level, these probabilities are influenced by the prevalence of disease in the study population and the cow-level test characteristics (Dohoo et al., 2009). At the herd-level, not only are these cow-level factors influential, but the population herd prevalence and the herd-level test characteristics are also important (Dohoo et al., 2009). However, as evidenced by our data, herd-level test characteristics vary based on within-herd prevalence and therefore, to use only 1 estimate of HSe for all herds would be inaccurate.

Another way to examine probability is using LR. This technique has been successfully used to predict the odds of MAP shedding with categorical individual cow ELISA data (Collins et al., 2005). In the context of these herd results, using LR allows for estimation of the likelihood of a herd being MAP-positive, given milk ELISA prevalence categories (Table 3.5). In herds with a milk ELISA prevalence between 0 and 2%, the test result was non-informative. However, when the milk ELISA prevalence was > 4%, within the limits of the available data from the study herds, test results in this category were associated with a MAP-positive herd.

3.5.1 Conclusion

No single HSe or HSp estimate for milk ELISA can be applied to all herds. In this study, with a range of within-herd prevalence values, HSe decreased significantly with

declining WHP. Conversely, whereas HSp was high overall, LR indicated that herds with less than 2% milk ELISA prevalence were likely to be false-positives nearly half the time. Finally, in addition to within-herd prevalence, the herd prevalence in the target population influenced interpretation of milk ELISA test results at the herd-level. The present results do not support the use of milk ELISA in low within-herd prevalence populations, but are consistent with previous consensus expert opinion suggesting milk ELISA testing was appropriate for control programs in high-prevalence herds.

3.6 References

Canadian Dairy Information Centre. 2013. Dairy in Canada at the farm 2013. Accessed November 18, 2013. http://dairyinfo.gc.ca/index_e.php?s1=dff-fcil&s2=farm-ferme&s3=nb.

Collins, M. T., V. Eggleston, and E. J. Manning. 2010. Successful control of Johne's disease in nine dairy herds: Results of a six-year field trial. J. Dairy Sci. 93:1638-1643.

Collins, M. T., I. A. Gardner, F. B. Garry, A. J. Roussel, and S. J. Wells. 2006. Consensus recommendations on diagnostic testing for the detection of paratuberculosis in cattle in the United States. J. Am. Vet. Med. Assoc. 229:1912-1919.

Collins, M. T., S. J. Wells, K. R. Petrini, J. E. Collins, R. D. Schultz, and R. H. Whitlock. 2005. Evaluation of five antibody detection tests for diagnosis of bovine paratuberculosis. Clin. Diagn. Lab. Immunol. 12:685-692.

DairyCo. 2012. Dairy statistics: An insider's guide 2012. Accessed November 18, 2013. http://www.dairyco.org.uk/non_umbraco/download.aspx?media=12943.

Dohoo, I. R., S. W. Martin, and H. Stryhn. 2009. Veterinary Epidemiologic Research. 2nd ed. VER, Inc., Charlottetown, P.E.I., Canada.

Gardner, I. A., S. S. Nielsen, R. J. Whittington, M. T. Collins, D. Bakker, B. Harris, S. Sreevatsan, J. E. Lombard, R. Sweeney, D. R. Smith, J. Gavalchin, and S. Eda. 2011. Consensus-based reporting standards for diagnostic test accuracy studies for paratuberculosis in ruminants. Prev. Vet. Med. 101:18-34.

Garry, F. 2011. Control of paratuberculosis in dairy herds. Vet. Clin. North Am. Food Anim. Pract. 27:599-607.

Hendrick, S., T. Duffield, K. Leslie, K. Lissemore, M. Archambault, and D. Kelton. 2005. The prevalence of milk and serum antibodies to Mycobacterium avium subspecies paratuberculosis in dairy herds in Ontario. Can. Vet. J. 46:1126-1129.

Kalis, C. H., M. T. Collins, H. W. Barkema, and J. W. Hesselink. 2004. Certification of herds as free of Mycobacterium paratuberculosis infection: Actual pooled faecal results versus certification model predictions. Prev. Vet. Med. 65:189-204.

Kalis, C. H., J. W. Hesselink, H. W. Barkema, and M. T. Collins. 2000. Culture of strategically pooled bovine fecal samples as a method to screen herds for paratuberculosis. J. Vet. Diagn. Invest. 12:547-551.

Klausen, J., A. Huda, L. Ekeroth, and P. Ahrens. 2003. Evaluation of serum and milk ELISAs for paratuberculosis in Danish dairy cattle. Prev. Vet. Med. 58:171-178.

Lavers, C., S.L.B. McKenna, I.R. Dohoo, H.W. Barkema, and G.P. Keefe. 2013. Evaluation of environmental fecal culture for Mycobacterium avium subspecies paratuberculosis detection in dairy herds and association with apparent within-herd prevalence. Can. Vet. J. 54:1053-1060.

Lombard, J. E. 2011. Epidemiology and economics of paratuberculosis. Vet. Clin. North Am. Food Anim. Pract. 27:525-535.

Lombard, J. E., T. M. Byrem, B. A. Wagner, and B. J. McCluskey. 2006. Comparison of milk and serum enzyme-linked immunosorbent assays for diagnosis of Mycobacterium avium subspecies paratuberculosis infection in dairy cattle. J. Vet. Diagn. Invest. 18:448-458.

Martin, S. W., M. Shoukri, and M. A. Thorburn. 1992. Evaluating the health status of herds based on tests applied to individuals. Prev. Vet. Med. 14:33-43.

McKenna, S. L., G. P. Keefe, H. W. Barkema, and D. C. Sockett. 2005. Evaluation of three ELISAs for Mycobacterium avium subsp. paratuberculosis using tissue and fecal culture as comparison standards. Vet. Microbiol. 110:105-111.

Nielsen, S. S. and N. Toft. 2008. Ante mortem diagnosis of paratuberculosis: A review of accuracies of ELISA, interferon-gamma assay and faecal culture techniques. Vet. Microbiol. 129:217-235.

Sockett, D. C., D. J. Carr, and M. T. Collins. 1992. Evaluation of conventional and radiometric fecal culture and a commercial DNA probe for diagnosis of Mycobacterium paratuberculosis infections in cattle. Can. J. Vet. Res. 56:148-153.

Tiwari, A., J. A. VanLeeuwen, S. L. McKenna, G. P. Keefe, and H. W. Barkema. 2006. Johne's disease in Canada part I: Clinical symptoms, pathophysiology, diagnosis, and prevalence in dairy herds. Can. Vet. J. 47:874-882.

USDA. 2010. Uniform Program Standards for the Voluntary Bovine Johne's Disease Control Program. USDA-APHIS-VS, National Animal Health Monitoring System, Fort Collins, CO (APHIS 91-45-016).

USDA. 2012. Farms, land in farms, and livestock operations. 2011 summary. Accessed November 18, 2013. http://usda01.library.cornell.edu/usda/nass/FarmLandIn//2010s/2012/FarmLandIn-02-17-2012.pdf

Wells, S. J., S. M. Godden, C. J. Lindeman, and J. E. Collins. 2003. Evaluation of bacteriologic culture of individual and pooled fecal samples for detection of Mycobacterium paratuberculosis in dairy cattle herds. J. Am. Vet. Med. Assoc. 223:1022-1025.

Wells, S. J., R. H. Whitlock, B. A. Wagner, J. Collins, F. Garry, H. Hirst, J. Lawrence, W. J. Saville, and A. L. Naugle. 2002. Sensitivity of test strategies used in the voluntary Johne's disease herd status program for detection of Mycobacterium paratuberculosis infection in dairy cattle herds. J. Am. Vet. Med. Assoc. 220:1053-1057.

Whitlock, R. H., S. J. Wells, R. W. Sweeney, and J. Van Tiem. 2000. ELISA and fecal culture for paratuberculosis (Johne's disease): Sensitivity and specificity of each method. Vet. Microbiol. 77:387-398.

Yokomizo, Y., H. Yugi, and R. Merkal. 1985. A method for avoiding false-positive reactions in an enzyme-linked immunosorbent assay (ELISA) for the diagnosis of bovine paratuberculosis. Jpn. J. Vet. Sci. 47:111-119.

Table 3.1. Definitions and acronyms used in an evaluation of 3 commercial milk ELISAs for detecting Mycobacterium avium subsp. paratuberculosis (MAP) in dairy herds.

Term	Definition
PFC	A pooled fecal sample consisting of individual fecal samples from 5 cows, based on cow-age
MAP-positive herd	A herd was classified as MAP-positive if any PFC during any test period was MAP culture-positive
MAP-negative herd	A herd was classified as MAP-negative if no PFC were MAP culture-positive. It is recognized that, even with 4 rounds of whole-herd testing, some MAP-positive low-prevalence herds may not have yielded any culture positive PFC. Nevertheless, these herds were treated as MAP-negative for the analyses
WHP	Cow fecal samples from positive PFC were cultured individually to estimate an apparent within-herd prevalence of MAP. Within-herd prevalence was calculated as the number of fecal culture-positive cows divided by the total number of cows from which fecal samples were collected. In the case of a positive PFC with all negative cow results, 1 cow in the positive PFC was considered positive.
Low-prevalence	MAP-positive herd with mean WHP $\leq$ 5%
High-prevalence	MAP-positive herd with mean WHP > 5%
Milk ELISA positive herd	If the proportion of cow-level ELISA positive test results in a herd test exceeded the specified cutoff, that herd was considered milk ELISA positive for the given test period.
Milk ELISA negative herd	If the proportion of cow-level ELISA positive test results in a herd test was below the specified cutoff, that herd was considered milk ELISA negative for the given test period.
Milk ELISA prevalence	Within-herd prevalence of milk ELISA test results
Herd prevalence	Proportion of MAP-positive herds

Table 3.2. Frequency of positive Mycobacterium avium subsp. paratuberculosis (MAP) cultures in 5 MAP-positive herds having both positive and negative pooled fecal culture herd test results during a 2 yr longitudinal study[a].

	No. cows fecal culture-positive/No. cows tested[b] (% positive)			
Herd	Round 1	Round 2	Round 3	Round 4
1	0/66	1/64 (1.6)	1/64 (1.6)	0/58
2	1/46 (2.2)	1/48 (2.1)	1/47 (2.1)	0/47
3	Not Tested	0/55	1/49 (2.0)	0/54
4	1/163 (0.6)	0/154	1/134 (0.7)	2/126 (1.6)
5	0/117	1/120 (0.8)	0/117	0/121

[a]The remaining 9 MAP-positive herds had positive PFC each time the herd was sampled. (Fig. 3.1 displays the range and mean MAP-prevalence for these 9 herds).
[b]Fecal samples from all cows tested were cultured in pools of 5; cow samples from positive pools were cultured individually.

Table 3.3. Distribution of Mycobacterium avium subsp. paratuberculosis (MAP) milk ELISA test results in 14 MAP-positive and 18 MAP-negative herds, during a 2 yr longitudinal study using 3 commercial milk ELISAs and 4 herd tests per herd.

No. ELISA positive cows per herd test	MAP-negative (18 herds)			MAP-positive (14 herds)[a]		
	No. (%) milk ELISA herd tests			No. (%) milk ELISA herd tests		
	ELISA A[b]	ELISA B[c]	ELISA C[d]	ELISA A	ELISA B	ELISA C
0	51 (70.8)	63 (87.5)	59 (81.9)	14 (28.6)	16 (32.7)	11 (22.4)
1	11 (15.3)	6 (8.3)	9 (12.5)	10 (20.4)	9 (18.4)	9 (18.4)
2	7 (9.7)	3 (4.2)	2 (2.8)	8 (16.3)	8 (16.3)	10 (20.4)
3-4	2 (2.3)	0 (0)	2 (2.8)	8 (16.3)	5 (10.2)	6 (12.2)
5-9	0 (0)	0 (0)	0 (0)	7 (14.3)	9 (18.4)	10 (20.4)
≥ 10	1 (1.4)	0 (0)	0 (0)	2 (4.1)	2 (4.1)	3 (6.1)

[a]Seven herds were added after the first round and were therefore only tested 3 times.
[b]Parachek2 Mycobacterium paratuberculosis test kit® (Prionics AG, Schlieren-Zurich, Switzerland).
[c]Mycobacterium paratuberculosis antibody test kit® (IDEXX, Westbrook, ME, USA).
[d]Paratuberculosis Indirect® (IDVet Innovative Diagnostics, Montpellier, France).

Table 3.4. Herd sensitivity and specificity for 3 commercial Mycobacterium avium subsp. paratuberculosis milk ELISA kits, using 3 cutoffs based on proportion of milk ELISA positive cows. Results obtained from pooled fecal culture during a 2 yr longitudinal study in 32 Canadian dairy herds were used as the herd reference standard.

	Herd Sensitivity (%) (95% CI)			Herd Specificity (%) (95% CI)		
	Milk ELISA Cutoff			Milk ELISA Cutoff		
ELISA	1%	2%	3%	1%	2%	3%
A[a]	64 (41-83)	59 (36-78)	45 (27-65)	79 (70-86)	80 (71-88)	93 (86-97)
B[b]	62 (36-83)	56 (32-77)	47 (27-68)	92 (84-96)	96 (89-98)	97 (90-99)
C[c]	72 (48-87)	63 (41-81)	53 (32-73)	88 (79-93)	92 (86-96)	96 (89-98)

[a]Parachek2 Mycobacterium paratuberculosis test kit® (Prionics AG, Schlieren-Zurich, Switzerland).
[b]Mycobacterium paratuberculosis antibody test kit® (IDEXX, Westbrook, ME, USA).
[c]Paratuberculosis Indirect® (IDVet Innovative Diagnostics, Montpellier, France).

Table 3.5. Likelihood ratios of Mycobacterium avium subsp. paratuberculosis (MAP) herd positivity for 4 categories of milk ELISA B within-herd prevalence.

Milk ELISA Prevalence (%)[a]	No. Herd Tests		Likelihood Ratio
	MAP-negative Herds	MAP-positive Herds[b]	
0	63	16	0.3
> 0 - ≤ 2	6	7	1.2
> 2 - ≤ 4	3	11	3.7
> 4	0	15	∞

[a]Herd milk ELISA B within-herd prevalence measurements were repeated every 6 mo during a 2 yr study period.
[b]A herd was considered MAP-positive if at least one pooled fecal culture was positive during the study period.

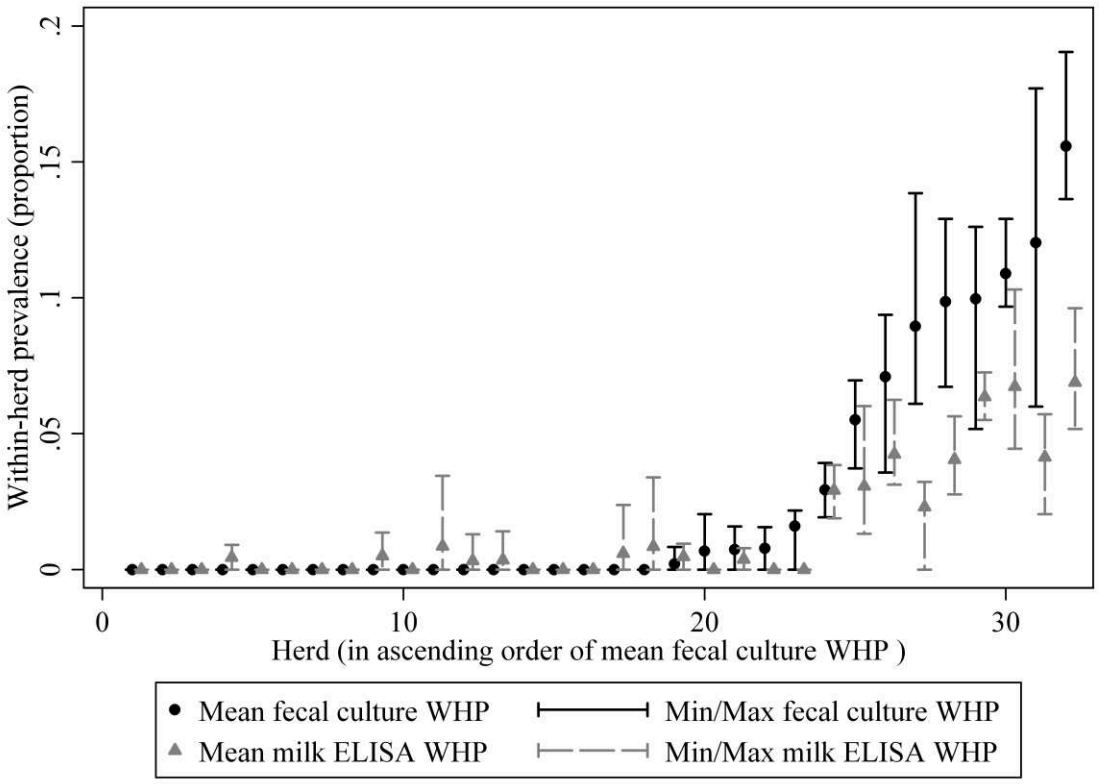

Figure 3.1. Mean, minimum and maximum within-herd Mycobacterium avium subsp. paratuberculosis prevalence (WHP) in 32 Maritime Canadian dairy herds using fecal culture and milk ELISA B (Mycobacterium paratuberculosis antibody test kit®).

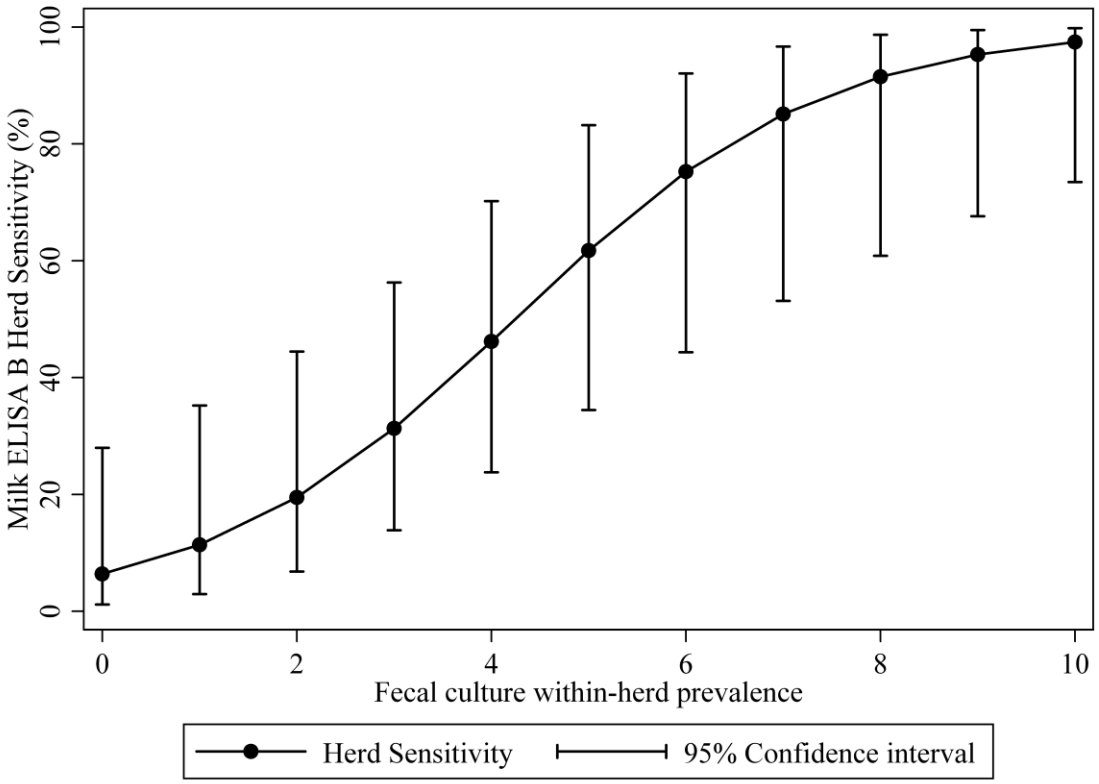

Figure 3.2. Estimated herd sensitivity of a commercial Mycobacterium avium subsp. paratuberculosis (MAP) milk ELISA (milk ELISA B; Mycobacterium paratuberculosis antibody test kit®), using 2% milk ELISA positive cows as the cutoff, with increasing fecal culture within-herd prevalence.

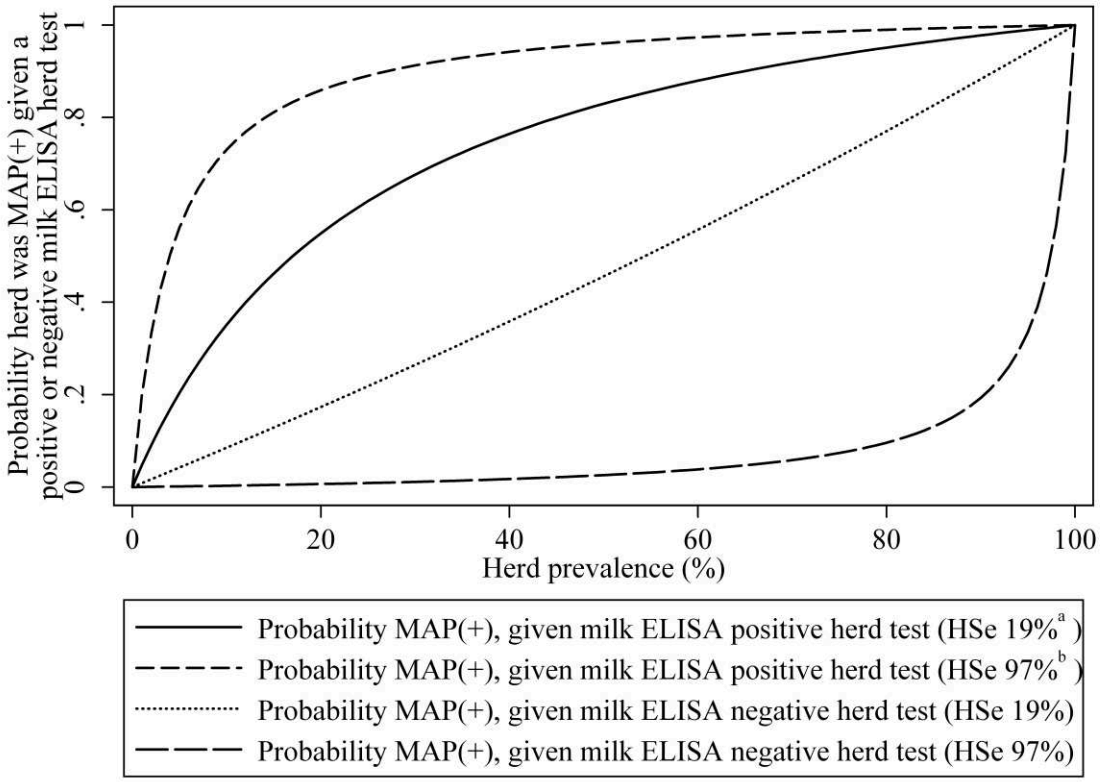

Figure 3.3. Influence of variable herd sensitivity and herd prevalence on the probability for a herd to be Mycobacterium avium subsp. paratuberculosis (MAP) positive, given a positive milk ELISA herd test result or a negative milk ELISA herd test result, using ELISA B (Mycobacterium paratuberculosis antibody test kit®) and 2% milk ELISA positive cows as the cutoff.

[a]HSe of 19% corresponds to the HSe estimate for a herd with a fecal culture within-herd prevalence (MAP-prevalence) of 2%.
[b]HSe of 97% corresponds to the HSe estimate for a herd with a MAP-prevalence of 10%.

CHAPTER 4

EVALUATION OF BETWEEN-TEST AGREEMENT FOR DIAGNOSIS OF BOVINE PARATUBERCULOSIS USING THREE COMMERCIAL MILK ELISAS

4.1 Abstract

This study evaluated between-test agreement on the cow-level diagnosis of Mycobacterium avium subspecies paratuberculosis for 3 commercial milk ELISA kits. Data were collected from 32 herds over a 2 yr period. Individual fecal and milk samples were collected from all lactating animals every 6 mo. Pooled fecal culture was conducted on samples pooled by cow-age, with 5 cow fecal samples in each pool. Subsequent individual fecal culture was performed on all cow fecal samples from positive pools. All fecal samples were cultured in a broth culture system, with confirmatory PCR performed on the broth of positive samples. Milk samples were analyzed using 3 commercial milk ELISA kits, and a cow-test was classified positive following the manufacturer recommended threshold for a positive cow result. Proportion agreed on negative was high: ELISA A/B: 0.992; ELISA A/C: 0.991; ELISA B/C: 0.996. Proportion agreed on positive was significantly lower: ELISA A/B: 0.61; ELISA A/C: 0.57; ELISA B/C: 0.79. Kappa statistics, indicating agreement beyond chance, ranged from moderate agreement to substantial agreement. Multi-level logistic regression models indicated that concurrent shedding of the bacterium increased the odds of agreement between test-positive milk ELISA results for all 3 paired test comparisons. Higher fecal culture within-herd prevalence increased the odds of agreement between test-positive milk ELISA results for ELISA A/B and ELISA A/C comparisons. Although the proportion of between-test agreement was lower for positive milk ELISA tests than for negative tests, the odds of between-test agreement on positive results increased when cows were shedding MAP.

4.2 Introduction

Paratuberculosis is caused by Mycobacterium avium subspecies paratuberculosis (**MAP**) (Chiodini et al., 1984), and has been reported in cattle worldwide (Harris and Barletta, 2001). The economic impact of the disease (Lombard et al., 2005; Nielsen et al., 2009), as well as a potential connection between the disease and Crohn's disease in humans (Barkema et al., 2011), make paratuberculosis very relevant to the dairy industry. The milk ELISA has become a popular cow-level test for MAP because results are obtained rapidly and at a low cost, relative to fecal culture (**FC**) (Tiwari et al., 2006). Estimates of cow-level sensitivity (**Se**) and specificity (**Sp**) of the milk ELISA are variable and depend on the stage of infection, with Se particularly noted to increase with progressive infection stages (Nielsen and Toft, 2008). Nielsen and Toft (2008) defined 3 stages of infection: infected (carrying MAP but not infectious); infectious (shedding MAP at the time of testing); and affected (clinical signs present). For infected animals, milk ELISA Se has been reported as 39% (Nielsen et al., 2002) and Sp as 96% (Nielsen et al., 2002) to 99.7% (Collins et al., 2005). For infectious animals, estimated Se of the milk ELISA has ranged from 21% (Lombard et al., 2006) to 61% (Hendrick et al., 2005) and Sp has ranged from 95% (Klausen et al., 2003; Hendrick et al., 2005) to 98% (Lombard et al., 2006). To the authors' knowledge, estimates for milk ELISA test characteristics in affected cows have not been published.

Milk ELISA test characteristics are often developed based on reference standard methods, with milk ELISA results being compared to a FC reference test. The evaluation of milk ELISA against FC has been suggested to be problematic for several reasons.

Although FC Sp approaches 100% (Nielsen and Toft, 2008), FC has potentially poor Se, with estimates ranging from 23% (McKenna et al., 2005) to 29% (Whitlock et al., 2000) for infected animals and 74% for infectious animals (Sockett et al., 1992). In addition to the potentially poor Se of FC as a reference test, FC and milk ELISA measure different aspects of the host response to MAP infection. Milk ELISA measures antibody response, which is typically detected after bacterial shedding is observed via FC, although antibody response has also been detected more than a year prior to detection of fecal shedding (Nielsen, 2008). Evaluating agreement between milk ELISAs allows for a comparison between two tests that are measuring the same host response. Given the relatively low Se of the milk ELISA relative to FC, it is important to determine if the milk ELISA test results at minimum agree with each other, and also if false-positive and false-negative results are consistent between kits. There have been a limited number of studies evaluating agreement between serum ELISA or serum and milk ELISA kits (Hendrick et al., 2005; McKenna et al., 2006; Lombard et al., 2006). However, agreement among commercial milk ELISA kits has not been evaluated.

Evaluating between-test agreement is also required because there is potential for producers and veterinarians to repeat a cow's milk ELISA test using a different ELISA kit, in order to confirm the initial test result. Given that there are several commercial milk ELISA kits available, it is important to evaluate if repeating the test, using the same sample but with a different kit, generates the same result. Practitioners and producers need to understand if this is a sound and beneficial practice. Therefore, the first objective of the present study was to quantify between-test agreement of commercial milk ELISA

kits at the cow-level, using milk ELISA data compiled from more than 4000 cows, within a study group comprising MAP-negative, low-prevalence and high-prevalence herds.

Although milk ELISA Se, relative to FC, has been noted to increase with progressive stages of MAP infection (Nielsen and Toft, 2008), and also with cow-age (Nielsen and Toft, 2006), it is not understood if agreement between milk ELISA tests is similarly affected by these cow-level factors. In addition, the impact of within-herd prevalence on between-test agreement is not known. Consensus recommendations indicate that the milk ELISA may be an appropriate cow-level test for control in herds with > 10% positive ELISA results (Collins et al., 2006). However, Johne's Disease is considered a disease of low to moderate within-herd prevalence and moderate herd prevalence (Lombard, 2011). As a result, many producers and veterinarians attempt to interpret cow-level milk ELISA results in herds below the consensus prevalence recommendation. Because these herds make up a large portion of the industry, it is important to evaluate the influence within-herd prevalence has on between-test agreement. Therefore, the second objective was to evaluate the herd-level and cow-level factors influencing positive milk ELISA between-test agreement.

4.3 Materials and Methods

4.3.1 Terminology

Cow fecal samples were cultured in pooled fecal cultures (**PFC**), with 5 cows per pool ordered by age (Kalis et al., 2000). Follow-up culture of all individual samples from positive PFC was used to establish cow-level FC results for each test. A cow was FC-

positive for a FC test if individual FC and confirmatory PCR were positive. If PFC was negative, a cow in that pool was considered FC-negative. It is recognized that there is no true gold standard diagnostic cow-test for paratuberculosis, and in particular, false-negative results are expected (Nielsen and Toft, 2008). However, for the purposes of the analyses, this was the definition of a FC-negative cow for each FC test.

A herd with no positive PFC during the entire study was classified MAP-negative. A herd was classified MAP-positive if, at any point during the study, a PFC was positive, indicating that MAP infectious cows were present in the herd. An apparent within-herd MAP-prevalence was calculated as the number of FC-positive cows divided by the number of cows from which fecal samples were collected during a herd test.

4.3.2 Selection of Herds

A total of 34 herds from the 3 Canadian Maritime provinces, Prince Edward Island, New Brunswick and Nova Scotia, participated in this 2 yr prospective study. Details on herd selection are provided in Lavers et al. (2013). Briefly, herd selection involved a non-random selection process based on prior risk assessments, in order to provide a mixture of MAP-negative, low-prevalence, and high-prevalence herds. In order for a herd's data to be included in the analyses, ≥ 3 rounds of individual cow FC had to be collected from a herd, and therefore 2 herds were excluded due to incomplete sampling. Median herd size was 66 milking cows (mean: 82; range: 28 to 220).

The MAP herd-status and MAP-prevalence for the 32 study herds is described in Lavers et al. (2013). Briefly, based on PFC results from the entire study period, 18 herds were MAP-negative and 14 herds were MAP-positive. The mean within-herd MAP-prevalence ranged from 0 to 15.6%, with a relatively even distribution of mean MAP-

prevalence values between the minimum and maximum value. Nine herds had a mean MAP-prevalence > 5% and 5 herds had a mean MAP-prevalence ≤ 5%.

4.3.3 Sample Collection

Individual fecal samples were collected from all lactating cows at 6 mo intervals during the 2 yr study period, using a clean rectal sleeve lubricated with sterile water. Approximately 30 g of feces was collected per rectum and placed in a clean, labeled 95 mL plastic specimen jar. Fecal samples were kept cool during transport back to the laboratory and were then frozen until time of processing. Fecal samples were frozen at -20˚C if processing occurred within 2 weeks and were frozen at -80˚C if processed between 2 and 6 weeks after collection. Individual milk samples were collected from lactating cows by Dairy Herd Improvement personnel (Valacta, Montreal, Quebec) at 6 mo intervals. Milk samples were preserved with bronopol and refrigerated at 4˚C. The aim was to have no more than 3 mo between a herd's milk and fecal sample collection.

4.3.4 Fecal Culture

All manure samples were processed by the technical laboratory staff at the Maritime Quality Milk Laboratory (Atlantic Veterinary College, Charlottetown, PEI, Canada), which was approved by USDA proficiency-testing for MAP FC. Fecal samples were processed and inoculated into ESP para-JEM® broth, according to the manufacturer's protocol, with the exception that samples were incubated for 49 d in the ESP® Culture System II (TREK Diagnostic Systems, Inc., Cleveland, OH, USA), rather than 42 d. The broth culture method, as well as confirmatory testing, has been described previously in Lavers et al. (2013). Briefly, after broth culture was complete, all broth

samples were examined microscopically for the presence of Mycobacteria using an acid-fast stain. Confirmatory PCR for detection of the hspX gene was performed on all samples positive by the culture system or microscopic visualization, using the VetAlert™ Johne's Real-Time PCR kit (Tetracore, Inc., Rockville, Maryland, USA) and following manufacturer's instructions.

4.3.5 Milk ELISAs

The three indirect ELISAs used were: ELISA A = Parachek2 Mycobacterium paratuberculosis test kit® (Prionics AG, Schlieren-Zurich, Switzerland); ELISA B = Mycobacterium paratuberculosis antibody test kit® (IDEXX, Westbrook, Maine, USA); and ELISA C = Paratuberculosis Indirect® (IDVet Innovative Diagnostics, Montpellier, France). ELISA A and B are licensed for the North American market. All three commercial milk ELISAs were performed by the technical laboratory staff at the Maritime Quality Milk Laboratory, which was USDA proficiency-tested for milk ELISA techniques. All 3 kits use an absorption step, which removes antibodies that cross-react with Mycobacterium phlei (Yokomizo et al., 1985), with the goal of minimizing cross-reactions with other Mycobacteria organisms. The milk ELISAs were performed following the manufacturers' instructions. The following milk ELISA procedure was used. All reagents and samples were brought to room temperature (18 - 25˚C) for at least 1 hr before use. Milk samples (150 µL ELISA A; 100 µL ELISA B; 80 µL ELISA C) and 10 µL each of positive and negative controls were added to the wells of the dilution plate. Diluent solution was then added to the wells containing milk samples (150 µL ELISA A; 100 µL ELISA B; 80 µL ELISA C) to create a 1:1 dilution of the milk sample. Diluent solution was added to the control wells (190 µL ELISA A and ELISA B; 110 µL ELISA

C) to create positive and negative control dilutions of 1:20 for ELISA A and B and 1:12 for ELISA C. The plates were gently shaken by hand and left to incubate at room temperature (30 min ELISA A; 15 min ELISA B and C). Following incubation, 100 µL of each sample and control was transferred by pipette to the coated ELISA microplates, and were then incubated for a further 45 min. At the end of this incubation period, the microplates were washed using the BioTek® Washer (BioTek Instruments, Inc., Winooski, VT, USA). ELISA A was washed 6 times, ELISA B was washed 3 times, with a 2 min soak on the third wash, and ELISA C was washed 3 times. After washing, the microplates were tapped dry onto clean absorbent paper, and 100 µL of diluted conjugate (1:100 dilution for ELISA A and B; 1:10 dilution for ELISA C) was added to each well. The microplate was incubated at room temperature for 30 min and then the wash was repeated. The washed microplate was tapped dry onto clean absorbent paper and then 100 µL of substrate solution was added to each well. ELISA A was read directly without incubation. For ELISA A, when the absorbance of the positive control was between 0.35 – 0.40 optical density (**OD**), measured at 630 nm using the BioTek® ELISA Reader, 50 µL of stop solution was added to each well of the ELISA A kit. ELISA B and C were incubated at room temperature in the dark (10 min ELISA B; 15 min ELISA C) after addition of the substrate solution, and then 100 µL of stop solution was added to each well. For all 3 ELISA kits, the absorbance of each well was read by the BioTek® ELISA Reader at 450 nm, following addition of the stop solutions to the wells.

ELISA A test output was reported as the mean negative control OD subtracted from the OD of the sample. ELISA B and C test output were reported as a sample to positive (S/P) ratio %, computed as: ((sample OD – negative control OD)/(positive

control OD – negative control OD)*100). A cow-level result was considered to be a test-positive milk ELISA result if the test result was greater than the threshold indicated in the kit instructions (ELISA A: 0.1 OD; ELISA B: 40% S/P%; ELISA C: 30% S/P%), and was otherwise considered a test-negative milk ELISA result.

4.3.6 Data Management and Statistical Analyses

Statistical analyses were performed using Stata/IC® Version 11.2 (StataCorp LP, College Station, Texas, USA). A P-value ≤ 0.05 was considered significant. The full dataset consisted of repeated cow test events for each of the 3 ELISAs, with a minimum of 1 test event and maximum of 4 test events per cow. Because it was not a study objective to evaluate agreement over time, for the analysis describing proportion agreement and kappa statistics, 1 test event per cow was randomly selected for the analysis. In order to be included in the random selection process, the cow test event had to include a milk sample collection from which all 3 commercial milk ELISAs being evaluated had complete results. To ensure the selection process was random, two subsequent analyses were conducted. First, additional random selection processes and subsequent analyses were repeated on the dataset to ensure results were similar, regardless of the cow test events randomly selected. Second, the analyses were repeated twice more, using the first, and subsequently the last, chronological cow test event to ensure results obtained were similar to the original random selection process.

Agreement statistics for ELISA A/B, ELISA A/C, and ELISA B/C between-test comparisons were calculated. These included overall observed proportion of agreement (Pa), as well as the proportion of paired milk ELISA test results that agreed on positive (Pa+), and negative (Pa-), test results (Cicchetti and Feinstein, 1990). Confidence

intervals (95%) for these proportions were calculated using the plus-four method (Agresti and Coull, 1998). The second agreement statistic calculated was Cohen's Kappa (kappa). This statistic is commonly used for subjective rating and represents the level of agreement beyond chance (Dohoo et al., 2009). In addition to kappa estimation, McNemar's test was also calculated in order to assess if proportions of positive results differed between ELISA results. A significant McNemar's test indicates disagreement between tests and reduces the interest in the kappa estimation (Dohoo et al., 2009). Due to the difference in scale for ELISA A/B and ELISA A/C comparisons, concordance correlation coefficients were not calculated.

In the second portion of the analysis, which evaluated factors influencing between-test agreement for test-positive milk ELISA results, 1 test event per cow was randomly selected, subject to 2 criteria. In order to allow for the evaluation of concurrent MAP shedding as a predictor of test-positive milk ELISA agreement, the first criterion was that a cow test event had to include complete results for both FC and all 3 commercial milk ELISA kits, with a maximum of 3 mo between fecal and milk sample collections. The second criterion was that at least 1 milk ELISA test result from the set of 3 results being compared was positive. Cow records meeting these 2 criteria were included in a random selection process, with 1 cow test event per cow being included in the analysis. Additional random subsets were selected, with subsequent analysis, to ensure that results were similar, regardless of the random selection process utilized.

Using this dataset, a multi-level logistic regression model was fitted for each of the 3 ELISA between-test comparisons (ELISA A/B; ELISA A/C; ELISA B/C). The outcome of the model was test-positive milk ELISA agreement between the 2 ELISA kits

being compared in the respective regression model. ("positive agreement" variable = 1). Therefore, when the milk ELISAs both reported negative results, or only 1 milk ELISA reported a positive result, the "positive agreement" variable = 0. The multi-level model structure accounted for clustering of test-positive milk ELISA cow results within herds. Predictors evaluated were mean MAP-prevalence (herd-level), cow-age (yrs) at time of testing (cow-level, categorical: < 3 yr, 3 - 4 yr, 4 - 5 yr, and > 5 yr), and concurrent cow MAP shedding (cow-level, dichotomous). The linear relationship between continuous predictors and the outcome was evaluated using fractional polynomial models, in order to determine if a transformation of the predictor provided a better fit for the model. Collinearity of variables and interaction terms were also evaluated. A P-value ≤ 0.05 was considered significant for inclusion in the final model.

4.4 Results

In the 32 study herds, 4569 cows had ≥ 1 milk or fecal samples collected during the project. There were 286 cows that had no milk samples collected, but had ≥ 1 fecal samples collected. From the 4283 cows contributing milk sample data, 516 (12%) cows had 4 milk ELISA test events, with complete results for all 3 commercial milk ELISAs. Similarly, 1199 (28%) cows had 3 test events, 1374 (32%) cows had 2 test events, and 1194 (28%) cows had 1 milk ELISA test event during the project. Mean herd cow-age at testing was 4.1 yrs (range: 2.9 to 5.5).

When 1 test event per cow was randomly selected, 2.1% (92/4283), 1.7% (73/4283), and 2.1% (89/4283) of cow results were positive for commercial milk ELISA

A, B, and C, respectively. The proportion of agreement was higher for negative results than positive results (Table 4.1). McNemar's test was significant for comparisons between ELISA A/B and ELISA B/ C, and non-significant for comparison between ELISA A/C. The highest level of agreement beyond chance occurred between ELISA B and C, with a kappa value of 0.79 (95% CI: 0.72 - 0.86) (Table 4.1). There were 2292 cow test events from MAP-negative herds, and 24 (1.1%), 8 (0.4%) and 9 (0.4%) of these cows had test-positive milk ELISA cow results for ELISA A, B and C, respectively. There were 1991 cow test events from MAP-positive herds. From these herds, there were 68 (3.4%), 65 (3.3%), and 80 (4.0%) cow test-positive milk ELISA results for ELISA A, B and C, respectively.

The quantitative ELISA results for the 3 pairs of ELISA comparisons are demonstrated graphically in Figures 4.1, 4.2, and 4.3. The vertical and horizontal reference lines indicate the manufacturer recommended threshold levels, and divide the scatterplot into four quadrants. The upper right and lower left quadrants indicate agreement, whereas the upper left and lower right quadrants indicate disagreement. For all 3 paired ELISA comparisons, the majority of the data points in the disagreement quadrants are relatively close to the threshold values. For all test kits, when results were discordant, between 86 - 100% of the ELISA values in the disagreement quadrants were below the median of the value for the corresponding positive ELISA.

Predictors of cow-level test-positive milk ELISA agreement in the multi-level logistic regression models are presented in Table 4.2. There were 130 cows from 23 herds with test-positive milk ELISA results in ELISA A/B model, 76 cows from 14 herds in ELISA A/C model, and 81 cows from 16 herds in ELISA A/C model. For all 3 models,

test-positive milk ELISA agreement was significantly more likely in cows that were concurrently shedding the MAP bacterium in their feces (OR ranging from 5.02 to 12.60), as detected by a positive concurrent FC result. For ELISA A/B and ELISA A/C models, test-positive milk ELISA agreement was positively associated with mean MAP-prevalence; for every 1% unit increase in mean MAP-prevalence, the odds of agreement increased by 1.15 (P-value: 0.009) and 1.17 (P-value: 0.005), respectively (Table 4.2). The mean MAP-prevalence did not significantly impact the odds of test-positive milk ELISA agreement for ELISA B/C. No transformation of the linear MAP-prevalence continuous variable was found to improve the model. Variables were not found to be significantly collinear, and interaction terms were not significant. Cow-age was not a significant predictor in the models. There was very little clustering within farms for all 3 pair comparisons (herd-level variances for ELISA A/B, B/C, and A/C were 0.001, 0.002, and 0.001, respectively).

4.5 Discussion

The present study evaluated the between-test agreement of 3 commercial milk ELISAs, and found that overall agreement was high, with kappa statistics indicating moderate to substantial agreement beyond that due to chance. However, the proportion of between-test agreement was lower for positive milk ELISA tests than for negative tests. The present study also evaluated the cow and herd-level factors influencing positive milk ELISA between-test agreement. Cow fecal shedding of the MAP bacterium and herd-

level MAP-prevalence were both significant predictors of agreement between positive milk ELISA test results.

A single measure of agreement, such as kappa or overall proportion of agreement, is an appealing method to summarize the results of an agreement study. However, a single descriptive statistic can be deceiving, and is not advised, if there is a symmetrical imbalance in the proportion of test-positive and test-negative results (Feinstein and Cicchetti, 1990). When this imbalance occurs, the effect of the smaller proportion is obscured within a single measure of overall agreement, and it is better to evaluate agreement by examining several statistics (Cicchetti and Feinstein, 1990). This symmetrical imbalance was apparent in the current data, where there was a substantially higher proportion of cows with test-negative milk ELISA results compared to cows with test-positive milk ELISA results. While agreement between negative test results was high, agreement between positive test results was significantly lower, indicating poorer agreement between different milk ELISA kits for test-positive results. A lower proportion of agreement in positive samples was also reported in a study evaluating agreement between serum ELISAs (McKenna et al., 2006). The lower agreement between positive results suggest that using a different milk ELISA kit to confirm a positive result is non-informative, as it will often yield a negative result.

Kappa statistics indicated moderate to substantial agreement between the 3 commercial milk ELISAs, beyond that expected due to chance. Kappa statistics between ELISA B/C indicated substantial agreement, and were significantly different from the moderate agreement estimated from kappa statistics for ELISA A/B and ELISA A/C comparisons. For ELISA A/B and B/C comparisons, there was a significant McNemar's

test, indicating a significant difference in the proportion of positive results between the tests and giving the kappa value less relevance (Caraguel et al., 2009). The kappa statistics reported here were similar to previous studies evaluating agreement between milk and serum ELISAs, which have reported fair (Lombard et al., 2006) and substantial agreement (Collins et al., 2005).

Kappa accuracy is influenced by extreme prevalence values (Dohoo et al., 2009), such as those observed with low-prevalence diseases like MAP, and therefore, kappa should be examined in study populations that are similar to the target population (Vach, 2005). In order to evaluate the agreement between commercial milk ELISAs that could be expected in the dairy industry, tested cows should represent all ranges of the infection spectrum, rather than targeted high-prevalence populations. Herds in this study were selected to be reflective of typical herds in the North American dairy industry, in order to allow for extrapolation to this target population.

Scatterplots provided a visual comparison of between-test agreement on a quantitative scale. These scatterplots illustrated that a large number of cows had negative results for both ELISAs, and also that, for many of the disagreeing cow-test results, the discordant positive results were relatively close to the specified ELISA kit threshold. A previous study evaluating serum ELISA agreement noted numerous pairs of discordant results where the test-positive milk ELISA result was much higher than the threshold (McKenna et al., 2006). It is possible that the ELISA kits used in the present study were more similar in terms of their antigen and conjugate components compared to the previous serum study.

For all 3 paired milk ELISA comparisons, fecal shedding of the MAP bacterium significantly increased the odds of test-positive milk ELISA agreement. This is similar to a previous longitudinal study, which reported that a greater proportion of cows consistently shedding the MAP bacterium had test-positive milk ELISA results (60 – 70% milk ELISA test-positive), as compared to cows intermittently or transiently shedding the bacterium (5 – 30% milk ELISA test-positive) (Nielsen, 2008). Test-positive milk ELISA agreement was less likely in cows that were not shedding the MAP bacterium. Antibody response and bacterial shedding do not necessarily begin at the same time, and longitudinal studies have indicated that a test-positive milk ELISA result can occur prior to, or following, detection of MAP shedding (Nielsen, 2008). Test-positive milk ELISA disagreement could occur because 1 ELISA kit is correctly identifying a MAP-infected cow that is not yet shedding the MAP bacterium, while the other ELISA kit is missing it. Conversely, because ELISA Sp is imperfect, the positive milk ELISA result in a discordant set may be a false-positive. Estimates of the milk ELISA Sp range from 95% (Hendrick et al., 2005) to just under 100% (Collins et al., 2005). Many of the discordant test-positive milk ELISA cow results are quite likely false-positive results reflecting the imperfect Sp.

The second predictor increasing the odds of test-positive milk ELISA agreement was the MAP-prevalence. This predictor was significant for test-positive agreement between ELISA A/B and ELISA A/C, but not for test-positive agreement between ELISA B/C. For ELISA A/B and ELISA A/C comparisons, respectively, the odds of test-positive milk ELISA agreement increased 15-17% with every 1% unit increase in MAP-prevalence. It is possible that MAP-prevalence was a significant predictor for ELISA A/B

and ELISA A/C agreement because of the higher proportion of ELISA A test-positive/ELISA B and C test-negative results observed in MAP-negative herds (0% MAP-prevalence). The higher proportion of disagreement in MAP-negative herds for ELISA A/B and ELISA A/C would contribute to the increased odds of test-positive milk ELISA agreement as the MAP-prevalence increased.

One limitation of the present study was the number of FC-positive and test-positive milk ELISA results. Even with 4000 cows in 32 herds, because paratuberculosis is a disease of low-prevalence, there are many more test-negative than test-positive results. Although a larger sample size is always desirable, to target only high-prevalence herds in order to increase the number of test-positive results creates limitations because these high-prevalence herds do not represent the majority of dairy herds, and outcomes may not be generalizable.

In this population of herds, cow fecal shedding of the MAP bacterium and higher herd MAP-prevalence both increased the odds of test-positive milk ELISA agreement, supporting consensus recommendations that the milk ELISA is best applied in herds with known infection, high MAP-prevalence, and clinical disease (Collins et al., 2006). Although overall and test-negative milk ELISA proportion of agreement was quite high, agreement on positive milk ELISA results was substantially lower. Unless there is a specified testing strategy in place for the application of 2 ELISA tests to 1 milk sample, such as an initial screening (higher sensitivity) followed by a repeat diagnostic (higher specificity) test (Dohoo et al, 2009), the interpretation of results from 2 milk ELISA kits may be confusing to interpret, based on the observed low proportions of positive agreement.

4.6 References

Agresti, A. and B. A. Coull. 1998. Approximate is better than 'exact' for interval estimation of binomial proportions. Am. Stat. 52:119-126.

Barkema, H. W., S. Hendrick, J. deBuck, S. Ghosh, G. G. Kaplan, and K. Rioux. 2011. Crohn's disease in humans and Johne's disease in cattle-Linked diseases? Pages 197-213 in Zoonotic Pathogens in the Food Chain. D. O. Krause and S. Hendrick ed. CAB International, Wallingford, UK.

Caraguel, C., H. Stryhn, N. Gagné, I. Dohoo, and L. Hammell. 2009. Traditional descriptive analysis and novel visual representation of diagnostic repeatability and reproducibility: Application to an infectious salmon anaemia virus RT-PCR assay. Prev. Vet. Med. 92:9-19.

Chiodini, R. J., H. J. Van Kruiningen, and R. S. Merkal. 1984. Ruminant paratuberculosis (Johne's disease): The current status and future prospects. Cornell Vet. 74:218-262.

Cicchetti, D. V. and A. R. Feinstein. 1990. High agreement but low kappa: II. Resolving the paradoxes. J. Clin. Epidemiol. 43:551-558.

Collins, M. T., I. A. Gardner, F. B. Garry, A. J. Roussel, and S. J. Wells. 2006. Consensus recommendations on diagnostic testing for the detection of paratuberculosis in cattle in the Uunited States. J. Am. Vet. Med. Assoc. 229:1912-1919.

Collins, M. T., S. J. Wells, K. R. Petrini, J. E. Collins, R. D. Schultz, and R. H. Whitlock. 2005. Evaluation of five antibody detection tests for diagnosis of bovine paratuberculosis. Clin. Diagn. Lab. Immunol. 12:685-692.

Dohoo, I. R., S. W. Martin, and H. Stryhn. 2009. Veterinary Epidemiologic Research. 2nd ed. VER, Inc., Charlottetown, P.E.I., Canada.

Feinstein, A. R. and D. V. Cicchetti. 1990. High agreement but low kappa: I. The problems of two paradoxes. J. Clin. Epidemiol. 43:543-549.

Harris, N. B. and R. G. Barletta. 2001. Mycobacterium avium subsp. paratuberculosis in veterinary medicine. Clin. Microbiol. Rev. 14:489-512.

Hendrick, S. H., T. E. Duffield, D. E. Kelton, K. E. Leslie, K. D. Lissemore, and M. Archambault. 2005. Evaluation of enzyme-linked immunosorbent assays performed on milk and serum samples for detection of paratuberculosis in lactating dairy cows. J. Am. Vet. Med. Assoc. 226:424-428.

Kalis, C. H., J. W. Hesselink, H. W. Barkema, and M. T. Collins. 2000. Culture of strategically pooled bovine fecal samples as a method to screen herds for paratuberculosis. J. Vet. Diagn. Invest. 12:547-551.

Klausen, J., A. Huda, L. Ekeroth, and P. Ahrens. 2003. Evaluation of serum and milk ELISAs for paratuberculosis in Danish dairy cattle. Prev. Vet. Med. 58:171-178.

Lavers, C., S.L.B. McKenna, I.R. Dohoo, H.W. Barkema, and G.P. Keefe. 2013. Evaluation of environmental fecal culture for Mycobacterium avium subspecies paratuberculosis detection in dairy herds and association with apparent within-herd prevalence. Can. Vet. J. 54:1053-1060.

Lombard, J. E. 2011. Epidemiology and economics of paratuberculosis. Vet. Clin. North Am. Food Anim. Pract. 27:525-535.

Lombard, J. E., T. M. Byrem, B. A. Wagner, and B. J. McCluskey. 2006. Comparison of milk and serum enzyme-linked immunosorbent assays for diagnosis of Mycobacterium avium subspecies paratuberculosis infection in dairy cattle. J. Vet. Diagn. Invest. 18:448-458.

Lombard, J. E., F. B. Garry, B. J. McCluskey, and B. A. Wagner. 2005. Risk of removal and effects on milk production associated with paratuberculosis status in dairy cows. J. Am. Vet. Med. Assoc. 227:1975-1981.

McKenna, S. L., H. W. Barkema, G. P. Keefe, and D. C. Sockett. 2006. Agreement between three ELISAs for Mycobacterium avium subsp. paratuberculosis in dairy cattle. Vet. Microbiol. 114:285-291.

McKenna, S. L., G. P. Keefe, H. W. Barkema, and D. C. Sockett. 2005. Evaluation of three ELISAs for Mycobacterium avium subsp. paratuberculosis using tissue and fecal culture as comparison standards. Vet. Microbiol. 110:105-111.

Nielsen, S. S. 2008. Transitions in diagnostic tests used for detection of Mycobacterium avium subsp. paratuberculosis infections in cattle. Vet. Microbiol. 132:274-282.

Nielsen, S. S. and N. Toft. 2008. Ante mortem diagnosis of paratuberculosis: A review of accuracies of ELISA, interferon-gamma assay and faecal culture techniques. Vet. Microbiol. 129:217-235.

Nielsen, S. S. and N. Toft. 2006. Age-specific characteristics of ELISA and fecal culture for purpose-specific testing for paratuberculosis. J. Dairy Sci. 89:569-579.

Nielsen, S. S., M. A. Krogh, and C. Enevoldsen. 2009. Time to the occurrence of a decline in milk production in cows with various paratuberculosis antibody profiles. J. Dairy Sci. 92:149-155.

Nielsen, S. S., C. Gronbaek, J. F. Agger, and H. Houe. 2002. Maximum-likelihood estimation of sensitivity and specificity of ELISAs and faecal culture for diagnosis of paratuberculosis. Prev. Vet. Med. 53:191-204.

Sockett, D. C., D. J. Carr, and M. T. Collins. 1992. Evaluation of conventional and radiometric fecal culture and a commercial DNA probe for diagnosis of Mycobacterium paratuberculosis infections in cattle. Can. J. Vet. Res. 56:148-153.

Tiwari, A., J. A. VanLeeuwen, S. L. McKenna, G. P. Keefe, and H. W. Barkema. 2006. Johne's disease in Canada part I: Clinical symptoms, pathophysiology, diagnosis, and prevalence in dairy herds. Can. Vet. J. 47:874-882.

Vach, W. 2005. The dependence of Cohen's kappa on the prevalence does not matter. J. Clin. Epidemiol. 58:655-661.

Whitlock, R. H., S. J. Wells, R. W. Sweeney, and J. Van Tiem. 2000. ELISA and fecal culture for paratuberculosis (Johne's disease): Sensitivity and specificity of each method. Vet. Microbiol. 77:387-398.

Yokomizo, Y., H. Yugi, and R. Merkal. 1985. A method for avoiding false-positive reactions in an enzyme-linked immunosorbent assay (ELISA) for the diagnosis of bovine paratuberculosis. Jpn. J. Vet. Sci. 47:111-119.

Table 4.1. Cross-classifications, proportions of agreement and kappa values for between-test agreement of cow-level milk ELISA Mycobacterium avium subsp. paratuberculosis test results using three commercial milk ELISA kits on 4283 cows from 32 herds.

ELISA Comparison	A[a]/B[b]	A/C[c]	B/C
Negative/Negative	4168	4154	4185
Negative/Positive	23	37	25
Positive/Negative	42	40	9
Positive/Positive	50	52	64
Proportion of agreement (Pa) (95% CI)	0.985 (0.981–0.988)	0.982 (0.978 – 0.986)	0.992 (0.989 – 0.994)
Proportion agreed on positive tests (Pa+) (95% CI)	0.61 (0.52 – 0.69)	0.57 (0.49 – 0.66)	0.79 (0.72 – 0.86)
Proportion agreed on negative tests (Pa-) (95% CI)	0.992 (0.990 – 0.994)	0.991 (0.989 – 0.993)	0.996 (0.995 – 0.997)
Kappa (95% CI)	0.60 (0.51-0.69)	0.57 (0.48-0.65)	0.79 (0.72-0.86)
Exact McNemar P-value	0.025	0.820	0.009

[a]Parachek2 Mycobacterium paratuberculosis test kit® (Prionics AG, Schlieren-Zurich, Switzerland).
[b]Mycobacterium paratuberculosis antibody test kit® (IDEXX, Westbrook, Maine, USA).
[c]Paratuberculosis Indirect® (IDVet Innovative Diagnostics, Montpellier, France).

Table 4.2. Results of multi-level logistic regression analysis for cow-level Mycobacterium avium subsp. paratuberculosis test-positive milk ELISA agreement for three commercial milk ELISA kits.

ELISA	Variable	Coefficient	Odds ratio (95% CI)	P-value
A[a]/B[b]	Intercept	-2.48		
	Concurrent FC result[d] (0=neg; 1=pos)	2.50	12.22 (4.67-31.98)	< 0.001
	Mean MAP-prevalence[e] (%)	0.14	1.15 (1.03-1.27)	0.009
	Variance (herd-level)	0.001		
A/C[c]	Intercept	-3.03		
	Concurrent FC result (0=neg; 1=pos)	2.53	12.60 (5.16-30.73)	< 0.001
	Mean MAP-prevalence (%)	0.16	1.17 (1.05-1.27)	0.005
	Variance (herd-level)	0.001		
B/C	Intercept	-0.15		
	Concurrent FC result (0=neg; 1=pos)	1.54	5.02 (1.99-12.67)	0.001
	Mean MAP-prevalence (%)	0.04	1.04 (0.94-1.14)	0.47
	Variance (herd-level)	0.002		

[a]Parachek2 Mycobacterium paratuberculosis test kit® (Prionics AG, Schlieren-Zurich, Switzerland).
[b]Mycobacterium paratuberculosis antibody test kit® (IDEXX, Westbrook, Maine, USA).
[c]Paratuberculosis Indirect® (IDVet Innovative Diagnostics, Montpellier, France).
[d]Pooled fecal culture with individual cow fecal culture of positive pools. A maximum of 90 d between milk and fecal sample collection.
[e]FC results were used to establish a mean apparent within-herd fecal culture prevalence estimate (MAP-prevalence), which was calculated as the number of FC-positive results divided by the total number of cows from which fecal samples were collected during the herd test.

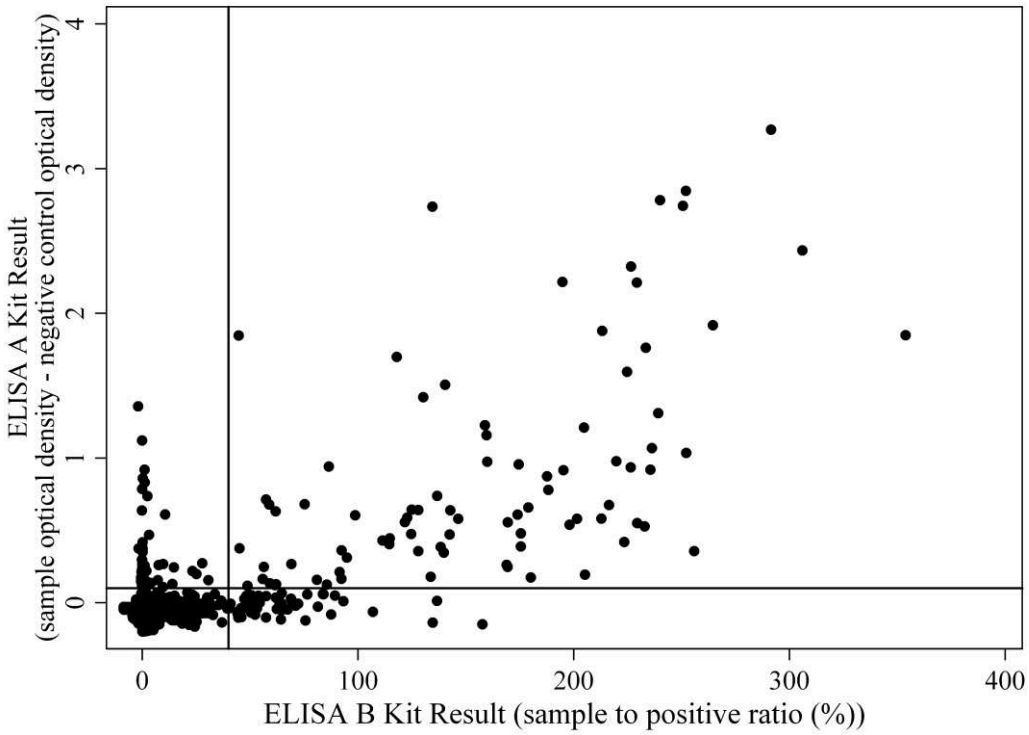

Figure 4.1. Scatterplot of cow-level milk ELISA results from 4283 cows in 32 herds for ELISA A (Parachek2 Mycobacterium paratuberculosis test kit®) and ELISA B (Mycobacterium paratuberculosis antibody test kit®) comparisons, overlaid with reference lines indicating the recommended threshold values for each ELISA kit.

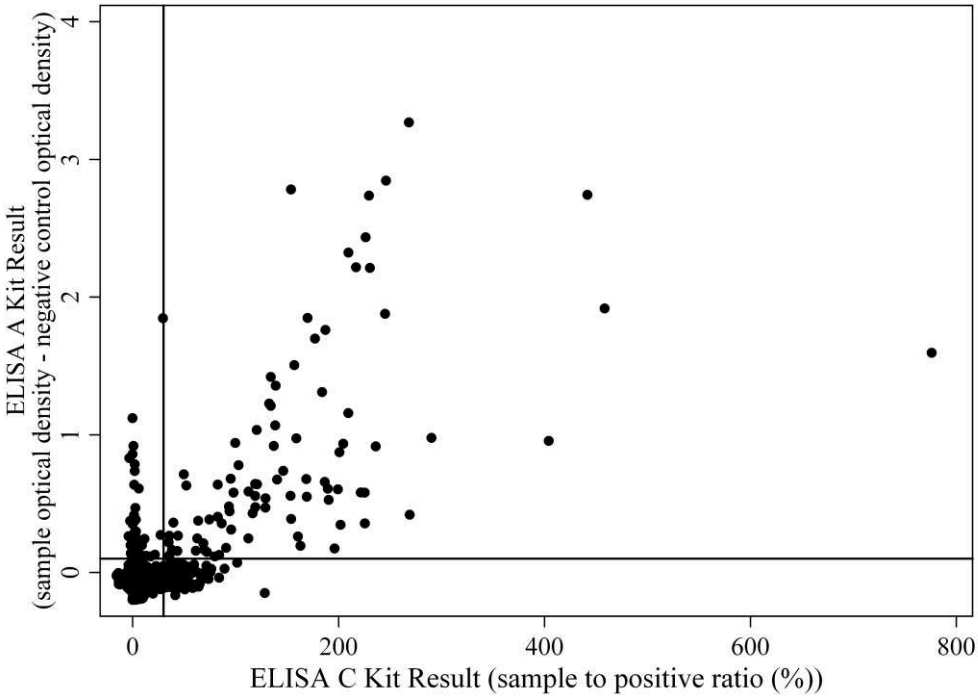

Figure 4.2. Scatterplot of cow-level milk ELISA results from 4283 cows in 32 herds for ELISA A (Parachek2 Mycobacterium paratuberculosis test kit®) and ELISA C (Paratuberculosis Indirect®) comparisons, overlaid with reference lines indicating the recommended threshold values for each ELISA kit.

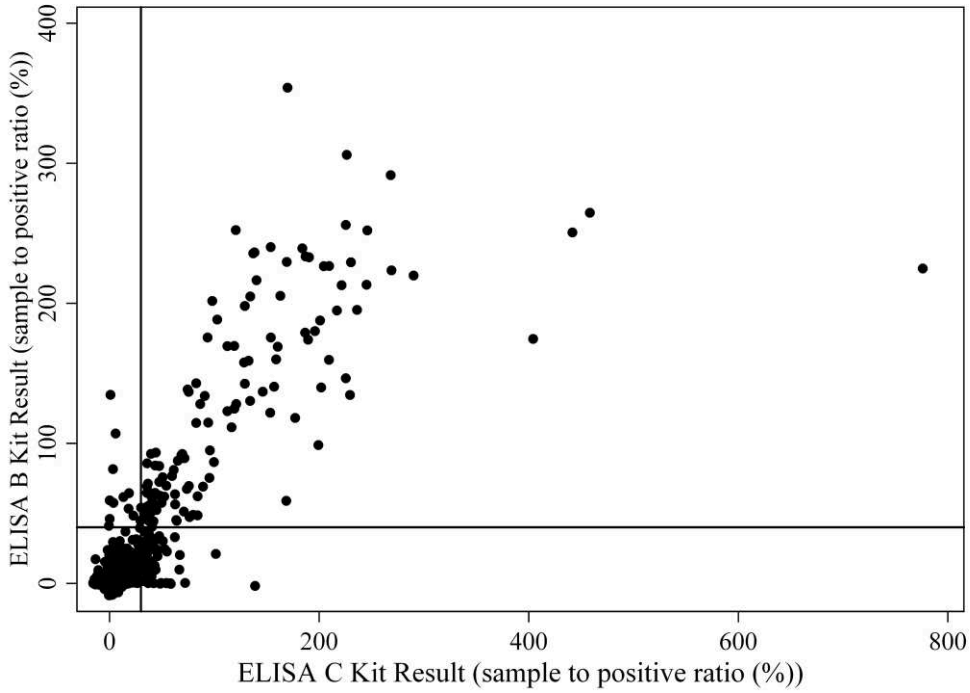

Figure 4.3. Scatterplot of cow-level milk ELISA results from 4283 cows in 32 herds for ELISA B (Mycobacterium paratuberculosis antibody test kit®) and ELISA C (Paratuberculosis Indirect®) comparisons, overlaid with reference lines indicating the recommended threshold values for each ELISA kit.

CHAPTER 5

TEST CHARACTERISTICS AND LIKELIHOOD RATIO INTERPRETATIONS FOR TWO COMMERCIAL BOVINE PARATUBERCULOSIS MILK ENZYME-LINKED IMMUNOSORBENT ASSAYS

5.1 Abstract

Test characteristics and categorical likelihood ratios were estimated for two commercial milk ELISAs using samples from 1829 cows in 15 Mycobacterium avium subspecies paratuberculosis (MAP) infected herds and 1889 cows from 17 non-infected herds. Pseudogold standard and latent class methods were used for test characteristic estimation; for pseudogold standards, a cow was considered infectious for MAP if a fecal sample collected within 3 mo of the milk sample was culture-positive using liquid culture. Sensitivity of ELISA A was 30.2% using a latent class model, and 28.4% using a pseudogold standard, while sensitivity of ELISA B was 34.6% using a latent class model, and 33.1% using a pseudogold standard. Specificity of ELISA A was 99.2% using a latent class model, and 99.5% using a pseudogold standard, while specificity of ELISA B was 99.4% using a latent class model, and 99.7% using a pseudogold standard. Using categorical likelihood analysis, the likelihood for a cow to be shedding MAP in its feces tended to increase with an increasing milk ELISA result. For ELISA A, a cow with an ELISA optical density < 0.05 was less likely to be shedding MAP in her feces than not shedding (Likelihood ratio = 0.7). Conversely, a cow with an ELISA optical density ≥ 0.50 was 196 times more likely to be shedding MAP in her feces than not shedding. Post-test probabilities for a cow to be MAP-infectious, given her pretest probability and categorical likelihood ratio, further extended the practical application of the data. Knowledge of the test characteristics of commercially available milk ELISAs are necessary when applying these kits in practice. Pseudogold statistical methods and latent class analyses produced similar estimates of milk ELISA test characteristics. Quantitative

interpretation of the milk ELISA result is an additional tool to provide further information from ELISA results within herd MAP control programs.

5.2 Introduction

Paratuberculosis (Johne's disease (**JD**)) is a chronic infectious enteritis of ruminants. The causative agent is a bacterium named Mycobacterium avium subspecies paratuberculosis (**MAP**) (Chiodini et al., 1984). Eradication of MAP is difficult, and successful control programs in infected herds should include both management changes that limit MAP exposure in calfhood, combined with MAP testing of adult cows. Preferably, testing strategies that identify the most infectious adult cows are used (Garry, 2011). Infectious cows should be detected and managed differently than their herd-mates, and even potentially removed from the herd, in order to facilitate MAP control and minimize impact of the organism on the herd (Nielsen and Toft, 2006). Due to limited sensitivity and imperfect specificity of tests, interpretation of diagnostic and screening test results is challenging. Test characteristics generally improve as cows progress through the 3 stages of disease, which are: infected (carrying MAP), infectious (infected and shedding MAP at the time of testing), and affected (occurrence of clinical signs in infectious animals) (Nielsen and Toft, 2008).

Detecting antibodies against MAP in milk, using ELISAs, is a test strategy used in many paratuberculosis control programs. Numerous milk ELISA kits are available, both commercial kits and in-house assays. Test characteristics of kits can vary substantially because kit components differ in terms of antigens and conjugates, and also

because different kits are optimized for different testing purposes (Nielsen, 2009). Because test characteristics are not transferable between test kits, it is very important to understand the test characteristics of the particular ELISA being applied to a herd (Nielsen, 2009). However, there are challenges associated with identifying the appropriate test characteristics of milk ELISAs. Firstly, many of the previously published test characteristics of milk ELISAs were based on in-house procedures (Nielsen et al., 2002; Klausen et al., 2003; Hendrick et al., 2005; Collins et al., 2005; Lombard et al., 2006), which have limited relevance to target populations outside those regions. Van Weering et al. (2007) evaluated a commercial ELISA kit that is widely available (Pourquier ELISA, Institut Pourquier, France); however, most of the test characteristics estimated in that study were relative to serum ELISA results only, which provides no indication as to how well the milk ELISA specifically detects infectious cows. In that study, the Pourquier milk ELISA was compared to fecal culture in a subset of 36 cows, all of which were moderate to high shedders, limiting the application of its results to moderate and high shedding cows only. In practice, a MAP-infected herd would be expected to contain cows exhibiting a wide range of shedding levels, from transient shedders to high shedders (Nielsen, 2008). As levels of MAP shedding increase, so does the probability for the milk ELISA to be positive (Nielsen, 2008). If the milk ELISA is only compared against fecal culture results obtained from cows shedding high levels of MAP, the estimated test characteristics will not be the same as those expected in a herd with a wider range of shedding levels.

In addition to limited use of commercially available milk ELISA kits in previous test evaluation publications, there is also a limited number of specificity estimates for

milk ELISA tests from non-infected herds. Paratuberculosis often has a low within-herd prevalence (Lombard, 2011), therefore test specificity is an important consideration, as imperfections in specificity translate into false-positive results. Depending on the herd's MAP control plan, the consequence of a false-positive test could be culling of a non-infected cow. With paratuberculosis, the best standard for estimation of specificity is to evaluate the test in herds that have had multiple negative tests of all adult cattle (Collins et al., 2005). Collins et al. (2005) evaluated an in-house milk ELISA on 7 uninfected herds (359 cows), and estimated a specificity of 99.7%. In the only published study on the specificity of a commercial milk ELISA kit in established non-infected herds, Van Weering et al. (2007) used data from 10 MAP-free herds (435 cows) to evaluate test characteristics of a commercial milk ELISA, and estimated a specificity of 100%, using the manufacturer's recommended cutoff.

Two statistical methods are commonly used to estimate test characteristics for assays where no strong reference standard exists, such as for paratuberculosis test evaluations. The first and more traditional approach is to use a pseudogold (reference) standard, where the case definition is based on available referent test(s), and consideration is given to possible misclassification bias in the interpretation of results. This is the approach used in many studies estimating milk ELISA test characteristics (Collins et al., 2005; Hendrick et al., 2005; Lombard et al., 2006; van Weering et al., 2007). The term reference standard implies an almost perfect test, which is not available for paratuberculosis (Nielsen and Toft, 2008), and for this reason, the term pseudogold standard was used in this chapter, in place of reference standard.

A second approach is the use of statistical methods to evaluate test characteristics in the absence of a gold standard reference test, commonly referred to as latent class analysis (Hui and Walter, 1980). Latent class analysis is performed through either maximum likelihood estimates (Hui and Walter, 1980) or Bayesian methodology (Joseph et al., 1995; Branscum et al., 2005). Although latent class analysis has not been used as commonly as pseudogold standard methods for the estimation of milk ELISA test characteristics in paratuberculosis diagnostic test evaluations (Nielsen et al., 2002; Nielsen and Toft, 2008), this approach seems appealing because the evaluated test is not compared to an imperfect reference, but rather the latent disease state of the cow. Norton et al. (2010) published estimates of serum ELISA test characteristics for paratuberculosis diagnosis using both pseudogold standard and Bayesian latent class models, and found that the 2 statistical methods produced similar estimates. Therefore, both statistical approaches hold merit for the estimation of milk ELISA test characteristics.

While test characteristics provide a means of determining the proportion of MAP infected cattle (or infectious or affected) detected by a test, producers and veterinarians also need to know how to interpret test results in order to make management and culling decisions on specific animals. The commercial milk ELISA kits produce a quantitative result, which is typically converted to a positive or negative diagnosis, based on the manufacturer's recommended cutoff. However, this dichotomization results in a loss of information, because the magnitude of the ELISA response correlates with MAP-infectious status (Nielsen, 2007). Categorical likelihood ratios (**LR**), similar to odds ratios, indicate how much more likely a particular category of test result is in infected versus non-infected animals (Blume, 2002; Deeks and Altman, 2004; Dohoo et al., 2009).

Likelihood ratios have been used to demonstrate a direct relationship between the magnitude of serum ELISA response and the odds of a cow shedding MAP in its feces (Collins et al., 2005). Applied to a paratuberculosis control program, categorical LRs can create a simple decision-making tool for producers and veterinarians based on the magnitude of the ELISA result (Collins, 2002; Naugle et al., 2003). However, reported categorical LR have been based only on serum ELISA data or in-house milk ELISAs (Collins, 2002; Collins et al., 2005), with a lack of published LR results for commercial kits applied to milk samples. Considering these factors, it would be useful and practical to estimate categorical LR that can be applied in the dairy industry, using commercial milk ELISA kits that are currently available.

The first objective of this research project was to determine sensitivity and specificity of 2 commercially available milk ELISA kits, using both pseudogold standard methods and latent class (no gold standard) models. The second objective was to evaluate if the quantitative ELISA result was associated with a cow being MAP-infectious.

5.3 Materials and Methods

5.3.1 Testing Purpose and Target Condition

The purpose for testing for this study was to detect cows shedding MAP, which is an important component of many paratuberculosis control programs (Nielsen and Toft, 2006; Garry, 2011). The target condition for analysis of sensitivity was a MAP-infectious cow, which is defined as a cow that is excreting MAP (as shown by a reference standard test) at the time of testing with the test under evaluation (Nielsen and Toft, 2008). A case

definition provides a practical definition of the target condition using a reference standard (Gardner et al., 2011). For pseudogold portions of the analysis, the case definition was a cow with a culture-positive fecal sample. The reference standard test (pseudogold standard) was pooled fecal culture, with subsequent culture of all individual samples from a positive pooled fecal culture, as has been described previously (Lavers et al., 2013b). There is no gold standard diagnostic test for paratuberculosis (Gardner et al., 2011), and given the low sensitivity of fecal culture in subclinical animals (Nielsen and Toft, 2008), it is recognized that a MAP-infectious cow in an infected herd could be misclassified as MAP non-infectious.

While sensitivity estimates to detect a MAP-infectious cow were developed from MAP-infected herds, specificity estimates were developed from a population of non-infected herds. A non-infected herd did not have any MAP culture-positive pooled feces or environmental manure samples collected during the 18 mo. study period. Conversely, a MAP-infected herd had 1 or more culture-positive pooled fecal sample or 1 or more culture positive environmental manure sample during the study period. The target condition for analyzing specificity was a MAP non-infected cow. The case definition was a cow that had never been fecal culture positive, and that came from a herd that had been repeatedly culture-negative on both cow and environmental manure samples. Collins et al. (2005) also estimated sensitivity and specificity for 2 separate target conditions, using similar populations to the current study. In a review of paratuberculosis diagnostic test evaluations, Nielsen and Toft (2008) described the target conditions from Collins et al. (2005) as MAP-infectious for sensitivity and MAP non-infected for specificity.

5.3.2 Herd Selection

Thirty-four herds from Prince Edward Island, New Brunswick and Nova Scotia, Canada, took part in this project, which ran from June 2009 to January 2011. Details on herd selection and demographics are provided in Lavers et al. (2013b). Briefly, herd selection was based on prior risk assessments, with the aim to enroll a combination of MAP non-infected, low-prevalence, and high-prevalence herds. A total of 27 herds were enrolled for the initial round of individual cow testing. Due to an inadequate number of MAP-infected herds on the initial round of testing, 7 additional herds were enrolled for the next 3 rounds of individual cow testing (cow feces and milk samples collected at 6, 12, and 18 mo) and the next 6 rounds of environmental sampling (manure samples collected at 6, 9, 12, 15, 18 and 21 mo). These 7 additional herds were selected because of 1 or more cases of paratuberculosis confirmed in the 2 yr previous to initiation of the research project. Four herds had cows with positive individual fecal culture, and 3 herds had cows with clinical signs suggestive of paratuberculosis and a positive ELISA test. Based on fecal culture, the highest mean within-herd MAP prevalence of study herds was 15.6%. Mean within-herd MAP prevalence was $\leq 5\%$ in 7 infected herds, and $> 5\%$ in the other 9 infected herds. Further information on fecal culture within-herd prevalence and milk ELISA within-herd prevalence has been published as another part of this research project (Lavers et al., 2013a).

The protocol for collection of cow fecal and environmental manure samples is described in Lavers et al. (2013b). The protocol for milk collection, completed by Dairy Herd Improvement personnel (Valacta, Montreal, Quebec), is described in Lavers et al. (2013a). Collection of milk and fecal samples from all lactating cows was done at 0, 6, 12, and 18 mo (4 rounds of testing). To be included in the statistical analyses, herds had

to have at least 3 rounds of individual cow fecal samples collected, resulting in 2 herds being excluded due to incomplete sampling.

5.3.3 Laboratory Processing

The fecal culture methodology used has been described previously in Lavers et al. (2013a). Briefly, individual cow fecal samples were pooled by age into pooled fecal cultures, with 5 cows in each pool (Kalis et al., 2000). Individual cow samples from positive pools were thawed and cultured individually to detect MAP-infectious cows. Fecal cultures were performed using the ESP® Culture System II (TREK Diagnostic Systems, Inc., Cleveland, OH, USA). As a confirmatory test, presumptive positive samples (acid-fast positive or positive culture system growth curves) were processed using VetAlert™ Johne's Real-Time PCR kit (Tetracore, Inc., Rockville, MD, USA).

Milk ELISA methodology has been described previously (Lavers et al., 2013a). The two commercial milk ELISAs used were milk ELISA A (Parachek2 Mycobacterium paratuberculosis test kit®, Prionics AG, Schlieren-Zurich, Switzerland) and milk ELISA B (Mycobacterium paratuberculosis antibody test kit®, IDEXX, Westbrook, ME, USA). Milk ELISAs were performed following the manufacturers' instructions. ELISA A test output was reported as the mean negative control optical density (**OD**) subtracted from the OD of the sample (also referred to as a corrected OD). ELISA B results were calculated as: (((sample OD – negative control OD)/(positive control OD – negative control OD))*100). Kit instructions describe ELISA B results as a sample to positive (**S/P**) ratio. Results described in these analyses reflect the S/P ratio described in the kit, but multiplied by a factor of 100, and are therefore referred to as S/P % in the current

study. A cow milk sample was defined as ELISA-positive based on the criterion provided in the relative kit.

5.3.4 Statistical Analyses

5.3.4.1 Data Management

Dairy cattle from 32 herds were included in the analyses. Cow-level tests were repeated over the study period in order to improve the confidence of the categorization of non-infected herds. At the cow-level, a reality of collecting repeated samples from commercial dairy herds was that, although some cows were sampled every 6 mo, other cows had missing data at some cow samplings due to their entry into or exit from the lactating herd during the study time frame. When cows had more than 1 set of fecal culture and milk ELISA test results (deemed a "complete set"), a random selection of complete sets of test results was selected to be included in the analysis. A second random selection process and subsequent analyses was conducted to ensure that results were similar, regardless of which cow results were randomly selected for the analyses. This random selection was deemed appropriate because the cow-level data were not designed or intended for use as a repeated measures analysis for this statistical portion of the research project.

5.3.4.2 Evaluation of Test Performance: Pseudogold Standard

Specificity of each milk ELISA was estimated using the test results from the 17 non-infected herds. Sensitivity of each milk ELISA was estimated from the 15 MAP-infected herds. A pseudogold standard was developed, where a cow was considered to be MAP-infectious if pooled fecal culture, and subsequent individual fecal culture, were

culture-positive. Test results for milk ELISA A and milk ELISA B were compared to the fecal culture result for estimation of test characteristics using a pseudogold standard.

5.3.4.3 Evaluation of Test Performance: Latent Class Model

Bayesian procedures using a Gibbs sampler in OpenBUGS (Medical Research Council and Imperial College, UK) were used to estimate test characteristics of milk ELISA A, milk ELISA B, and fecal culture. Populations 1 through 15, inclusive, were MAP-infected herds. Each herd was considered a distinct population because prevalence was not expected to be constant across all herds. The other 17 study herds were MAP non-infected, and were combined into 1 population (population 16), because prevalence was expected to be the same across herds. Test accuracy was assumed to be constant across populations. Informed priors based on expert opinion were used for specificity of fecal culture and milk ELISA (Norton et al., 2010) and sensitivity of fecal culture and milk ELISA (Collins et al., 2006), as presented in Table 5.1. The informed prior for the population prevalence was based on repeated testing of the study herds published previously (Lavers et al., 2013a; Lavers et al., 2013b). Using these priors, Beta distributions were determined using the BetaBuster 1.0 software package (UC Davis Graduate Group in Epidemiology, Davis, CA) by specifying the expected mode and expected 95^{th} percentile (Table 5.1).

When tests have a similar biological basis, estimates of test accuracy that do not account for test dependence can be misleading (Branscum et al., 2005). Bayesian models have been generalized to allow for test dependence (Georgiadis et al., 2003). Conditional dependence between tests was evaluated by comparing 4 models: (1) conditional independence between all tests; (2) conditional dependence between fecal culture and

milk ELISA A; (3) conditional dependence between fecal culture and milk ELISA B; and (4) conditional dependence between milk ELISA A and milk ELISA B. Two versions of Bayesian models were created: with either fixed or random effects for the infected herds. Modeling herds by random effects allowed for quantification of between-herd variation in the prevalence in the infected herds (Dohoo et al., 2009). The models with herds as random and fixed effects utilized the same code, data, and prior information for all other model parameters. Herd random effects were assumed normally distributed on logistic scale. The prior distribution for the precision (inverse variance) was a gamma distribution (0.001, 0.001). The random effects model equation on logistic scale included an intercept whose prior distribution was determined from the beta distribution prior for herd prevalence.

Latent class models were run for 100,000 iterations after discarding an initial burn-in of 10,000 iterations. Three chains with distinct initial values were used in order to assess convergence to a posterior distribution, which was evaluated by monitoring the Brooks-Gelman-Rubin plots. The deviance information criterion and Bayesian P-values of models were monitored to assess model fit. The mean and 95% probability intervals were reported for distributions of interest.

Sensitivity analysis was performed to ensure repeatability of results. The model was re-run after a single informed prior was changed within a biologically sensible limit. This procedure was performed after altering priors for each of the following: sensitivity of milk ELISA A, milk ELISA B, and fecal culture; specificity of milk ELISA A, milk ELISA B; and fecal culture; within-herd prevalence of MAP infection for both the MAP-infected and non-infected herds.

5.3.4.4 Categorical Likelihood Ratios

Categorical likelihood ratios (**LR**) were developed for milk ELISA A and milk ELISA B. In order to calculate the LR within each category, the diseased population was comprised of cows with a positive fecal culture result (MAP-infectious cows), and the non-diseased population consisted of all cows from non-infected herds. Creation of these case and control populations is similar to previous methods published evaluating LR for a serum ELISA (Collins, 2002).

No publications or industry-applied categories were found for LR analysis using either ELISA A or ELISA B on milk samples. As a result, 5 LR categories similar to those published by Collins (2002), where 2 categories were below the manufacturer cutoff and 3 were above were defined. Development of categories was based initially on publications for similar serum ELISA kits or in-house milk ELISA kits. These categories were then modified using an approach similar to Collins (2002); consideration was given to the magnitude of the ELISA results, as well as clinical experience, and the objective to create practical categories for application in practice. Categories for ELISA A were based on a study of an in-house milk ELISA (Collins et al., 2005), which had the same manufacturer cutoff as ELISA A (≥ 0.10). Categories for ELISA B LR were developed using categories similar to those used by Collins (2002) when evaluating serum IDEXX® ELISA LR.

Likelihood ratios were calculated for each ELISA output category as the probability of the result in the diseased population, divided by the probability of the result in the non-diseased population (Dohoo et al., 2009). Confidence intervals (95%) for LR were calculated using the same method applied for risk ratios (Deeks and Altman, 2004),

which incorporates a logarithmic transformation of the LR, which allows for the estimation to follow a normal distribution, and subsequent calculation of confidence intervals (Dohoo et al., 2009). Using the categorical LR estimates for ELISA A and pre-test probability of a cow to be MAP-infectious (within-herd MAP prevalence), the post-test probabilities for MAP-infectiousness were calculated (Dohoo et al., 2009).

5.4 Results

Median herd size was 66 milking cows (mean: 82; range: 28 to 220). Amongst all cows tested from all herds, the median age at testing was 4.0 yr (mean: 4.4 yr; range: 1.8 – 17.3 yr).

5.4.1 Estimation of Test Characteristics

For the pseudogold test evaluation, Table 5.2 provides a cross-tabulation of fecal culture, milk ELISA A and milk ELISA B cow-level test results, from 1889 cows within 17 non-infected herds and 1829 cows within 15 infected herds. Table 5.3 provides the point estimates for test characteristics for the pseudogold standard.

For latent class modeling of test characteristics, a model with conditional dependence between the 2 milk ELISAs was selected, because it had the greatest covariance between two tests. There was no substantial difference between the latent class models with herd as a fixed or a random effect, and the fixed effects model was chosen for its fewer model assumptions.

Using latent class models, the sensitivity of ELISA A and ELISA B were 28.4% and 33.1%, respectively. Using the latent class models, specificity of milk ELISA A and ELISA B were 99.3 and 99.7%, respectively (Table 5.3). The point estimates for test characteristics between the latent class and pseudogold standard were not substantially different.

5.4.2 Likelihood Ratios

Categorical LR, based on 5 categories of ELISA A results, are presented in Table 5.4. Cows with milk ELISA A OD values ≤ 0.05 were 0.7 times as likely to be MAP-infectious as non-infected cows. Cows with OD values ≥ 0.50 were 196 times more likely to be infectious than non-infected.

Categorical LR for ELISA B based on 5 S/P % categories are displayed in Table 5.5. Cows with ELISA B S/P % ≤ 20 were 0.6 times as likely to be MAP-infectious as non-infected. Cows with S/P % ≥ 100 were 465 times more likely to be infectious than non-infected.

Table 5.6 provides post-test probabilities for a cow to be MAP-infectious, given her pretest probability and ELISA A categorical LR. For example, a cow from a herd with a 2% MAP within-herd prevalence and a positive ELISA A reading between 0.100 and 0.299 had a 28% probability of being MAP-infectious. For a cow from a herd with a 30% within-herd prevalence and an ELISA A reading in the same range, this probability was 89%. These post-test probabilities were not calculated for ELISA B, as ELISA B LR in the 3 middle categories largely overlapped, precluding further extension to post-test probabilities.

5.5 Discussion

Paratuberculosis is considered a disease of low to moderate within-herd prevalence (Lombard et al., 2011). Therefore, milk ELISA tests are frequently applied to non-infected cows, and accurate specificity estimates of commercial paratuberculosis ELISA kits are vital, because of the potential for false-positive results. An accurate specificity is particularly important when a positive test can result in culling of the animal. Specificity of paratuberculosis tests can be difficult to determine because the prolonged latent period of the disease prevents the accurate use of a reference test to establish absence of infection in a cow. A negative fecal culture from a cow in a MAP-infected herd is not adequate to define the cow as non-infected (Collins et al., 2005). Instead, the best method to define absence of infection in a cow is to determine whether the herd within which she was raised is infected or not. In the current study, repeated cow-level (pooled fecal culture) and herd-level (environmental culture) testing allowed for documentation of a large population of non-infected herds. From this group of 17 non-infected herds, specificity estimates of milk ELISAs were calculated.

Specificity estimates for the 2 commercial milk ELISAs were not substantially different from each other. Specificity of ELISA A using a pseudogold standard was 99.5%, while it was 99.3% using a latent class model. Specificity of ELISA B was 99.7% for both statistical approaches. Published estimates of specificity have been lower, at 95% (Hendrick et al., 2005), 96% (Nielsen et al., 2002), and 98% (Lombard et al., 2006). However, these specificity estimates were estimated in infected herds. The prolonged

latent stage of paratuberculosis precludes the definitive determination that a cow from an infected herd is a non-infected cow, based on a negative fecal culture result. Establishing that a cow is non-infected must therefore be based on demonstration of absence of infection for the entire herd from which the cow was raised (Collins et al., 2005).Two studies have evaluated the specificity of the milk ELISA in herds that were repeatedly test negative, with estimates of 99.7% using an in-house ELISA (Collins et al., 2005), and 100% using the Pourquier ELISA at the manufacturer recommended cutoff (van Weering et al., 2007). The specificity estimates from the current data for 2 commercial milk ELISAs were very similar to these estimates. However, the current estimates of specificity have several advantages over previous work. First, the specificity estimates came from a large population of 1889 cows. Additionally, 2 statistical methods were used: the more traditional method of comparing the milk ELISA result to a pseudogold standard, as well as a latent class model that allows for an assumption of no gold standard test. These 2 statistical methods produced estimates that were not substantially different. Finally, 2 commercial milk ELISA kits were used, following manufacturer recommended cut-offs, making results practical for application to the dairy industry.

Sensitivity estimates for detection of MAP-infectious cows for 2 commercial milk ELISAs were not substantially different from each other, and ranged from 28.4 to 34.6%. Previous estimates of sensitivity of the milk ELISA for detection of MAP-infectious cows, using in-house ELISAs, have included 21% (Lombard et al., 2006), 29% (Collins et al., 2005), 54% (Klausen et al., 2003), and 61% (Hendrick et al., 2005). An evaluation of milk ELISA test characteristics using a commercial kit (van Weering et al., 2007) only estimated sensitivity of the milk ELISA in moderate to high fecal shedders. Comparing

milk ELISA test results to this subset of infectious cows, van Weering et al. (2007) estimated an overall sensitivity of 89%, and a sensitivity of 97% when only high shedders were included. As a comparison group, moderate and high shedding cows produce an upward biased estimate of sensitivity. In practical situations, it would be expected that a MAP-infected herd would have a mixture of non-shedders, transient, intermittent, moderate and high shedders (Nielsen, 2008). The high milk ELISA sensitivities reported by van Weering et al. (2007), relative to previously published estimates, are very likely reflective of the artificial population of moderate and high MAP shedders within which the milk ELISA was evaluated. This variability in sensitivity results underscores the importance of evaluating the ELISA within a study population similar to the target population. It is expected that most MAP-infected herds have a low to moderate within-herd prevalence (Lombard, 2011), which is reflected in the current study population. The current study sensitivity estimates are not substantially different to estimates of Lombard et al. (2006), which also had a mixture of low and moderate within-herd prevalence. Seven of the 18 herds used for sensitivity estimates in the Lombard et al. (2006) analysis had a fecal culture within-herd prevalence > 10%, with the remaining 11 herds having a prevalence < 10%. Collins et al. (2005) had higher within-herd fecal culture prevalence herds than the current study (7 herds, ranging from 9.5-32.5%), although the sensitivity estimates were similar to the current study. The higher sensitivity estimates from Hendrick et al. (2005) and Klausen et al. (2003) could reflect the fact that Hendrick et al. (2005) used 9 known infected herds (prevalence estimates not available) and Klausen et al. (2003) used 6 herds that had a history of several clinical JD cases each year, which is indicative of high MAP prevalence.

Despite the use of both repeated cow-level and herd-level testing in the current study to establish a non-infected status for a herd, it is possible that a MAP-infected herd was misclassified as non-infected. Low-prevalence herds are most susceptible to misclassification as a non-infected herd (Kalis et al., 2004). Although pooled fecal samples have a limited loss in test sensitivity relative to individual fecal samples (Kalis et al., 2000; Wells et al., 2003; Van Schaik et al., 2003; Collins et al., 2006), there is no perfect 'gold standard' test for MAP at the herd-level (Gardner et al., 2011) from which to establish a herd's MAP status. Repeated testing was used to minimize risk of misclassification and any potential impact on test evaluation outcomes, but any paratuberculosis study based on diagnostic test outcomes should be interpreted with some caution (Hendrick et al., 2006; Nielsen and Toft, 2008). An alternative to a case definition based on pooled fecal culture would have been a case definition developed from individual culture of fecal samples, but this would have been cost-prohibitive for this study.

Although estimation of fecal culture test characteristics was not an objective of the research, they were estimated as part of the latent class model. The latent class model estimated a sensitivity of 78% and a specificity of 99.96%. Published specificity estimates for fecal culture have ranged from 98 to 100% (Nielsen and Toft, 2008). Sensitivity estimates have been variable, ranging from 23% (McKenna et al., 2005) to 29% (Whitlock et al., 2000) for infected cattle and 74% for infectious animals (Sockett et al., 1992). The sensitivity estimate for fecal culture in the current study, designed to detect infectious animals, is similar to the estimate of Sockett et al. (1992). It is also close

to the upper range of fecal culture sensitivities (75%) estimated by Norton et al. (2010) with latent class models.

There was no substantial difference between milk ELISA test characteristic estimates using pseudogold standards or latent class models. This finding is similar to Norton et al. (2010), where serum ELISA test characteristics were evaluated using latent class models, with a pseudogold standard for comparison, and estimates were not substantially different between the 2 methods. In the current study, results of both statistical methods are presented because each produces relevant estimates for comparison purposes. The pseudogold standard allows for comparison of results between the current study and previous publications using this method of statistical evaluation (Klausen et al., 2003; Collins et al., 2005; Hendrick et al., 2005; Lombard et al., 2006; van Weering et al., 2007). However, estimates determined using the more traditional pseudogold standards have been questioned because they assume an accurate reference test, which is challenging in paratuberculosis research (Nielsen et al., 2002; Wang et al., 2011). The latent class model allows for estimation of test characteristics without requiring a decision regarding the referent disease status of the animal. Because both methods produced similar estimates, the pseudogold standard was preferred for these data, as it achieved the same result as a computationally more complex method, but is more comprehensible for a wider audience. For these reasons, the LR analysis was based on the pseudogold standard classification of cows as infectious and on the classification of herds (and cows within these herds) as non-infected by repeated fecal sample monitoring.

Likelihood ratios are a statistical tool that has the potential to extend diagnostic information from ELISA results, by considering the magnitude of the result, rather than a dichotomized diagnosis. Likelihood ratios for paratuberculosis have generally been estimated using a traditionally accepted approach (Collins, 2002; Naugle et al., 2003; Collins et al., 2005; Dohoo et al., 2009), although there have been LR analyses developed using Bayesian methodology (Fosgate et al., 2006). In the current analysis, the pseudogold standard approach to LR analysis (Dohoo et al., 2009) was selected for 2 reasons. First, data were sparse in the middle categories of ELISA values, and even with the methods used in the current analysis, confidence intervals overlapped. It was expected that probability intervals would be larger for a latent class analysis than confidence intervals surrounding pseudogold standard estimates, as the model would be accounting for uncertainty regarding the latent disease state of the cow. Secondly, estimates of test characteristics were not substantially different for the pseudogold standard and Bayesian models, suggesting there would be little gain in development of Bayesian LR.

Categorical LRs of both ELISA A and B for a cow being MAP-infectious tended to increase with an increasing OD value. A similar relationship was reported for an in-house milk ELISA (Collins et al., 2005), ELISA A applied to serum (Collins et al., 2005), as well as the ELISA B kit used on serum samples (Collins, 2002; Naugle et al., 2003). The lowest and highest categories of ELISA B were similar to ELISA A, with the lowest category having a very low LR (0.6), and the highest category a very high LR (465). However, the LRs of the 3 middle categories of ELISA B were all similar. Likelihood ratios above 10 are considered strong evidence in support of disease diagnosis (Deeks and Altman, 2004), and all 3 of these categories were higher than 10. Therefore, based on

LR these 3 categories were all supportive evidence of a cow being MAP-infectious. However, it was not possible to distinguish these 3 categories further. Collins (2002) did not report confidence intervals around LR estimates, but did note sparse data in the middle categories that limited the precision of LR estimates. The same limitation was noted in the current LR estimations for both ELISA A and B. Potentially, if this study was repeated with a larger sample size, a more clear and distinct direct relationship between ELISA B categories and probability of infectiousness could be seen. A second possibility for the lack of distinction between categories in ELISA B is related to the manufacturer cutoff. Using ELISA B, Van Weering et al. (2007) found that a change in cutoff from the manufacturer's recommendation (30%) to a 20% S/P % resulted in an increase in relative sensitivity by 7 percentage points, with only a small decrease in specificity, from 100% to 99.8%. In the current dataset, as a comparison, if ELISA B cutoff was changed to 20% S/P %, using pseudogold standards, the sensitivity increased from 34.6 to 42.3%, with a small decrease in specificity, from 99.8 to 99.7%. It was not an objective of this study to investigate the impact on test characteristics when kit cutoffs were altered, but rather help guide the clinical application of commercial kits using manufacturer guidelines. However, this represents a potential objective that should be investigated in a future study.

One of the advantages of the LR is that it allows for the use of quantitative results, rather than limiting the results to a dichotomous diagnostic outcome (Fosgate et al., 2006). For example, Collins (2002) applied LR to 5 categories of ELISA output that translated into 5 interpretations of the ELISA outcome: 'negative', 'suspect', 'weak positive', 'positive' and 'strong positive', and provided some explanation and

recommendation for how these interpretations would work within a MAP control program. This is an example of a practical application of the data for the dairy industry.

Post-test probabilities that a cow was MAP-infectious were calculated using the LRs and estimated within-herd MAP prevalence (pre-test probability of MAP infection). The post-test probabilities are useful as tools in decision analysis models, as they can inform economic outcomes based on the magnitude of the ELISA result (Collins et al., 2005). Table 5.6 provides example calculations of these post-test probabilities. If, for example, a cow has an ELISA result in the upper category of the ELISA output, its likelihood of being MAP-infectious is very high, and a confirmatory test such as fecal culture would provide very little additional information regarding her MAP disease status. The large confidence intervals around the LR estimates induce substantial uncertainty in these post-test probabilities. Even so, they are a valuable example of the influence that both within-herd prevalence and magnitude of ELISA result have on the interpretation of a cow-level result.

The establishment of test characteristics for 2 commercial milk ELISAs provides valuable information for practitioners and producers using these kits in practical applications in the dairy industry. Estimated within-herd prevalence and LR allow for a more precise estimation that a cow is MAP-infectious. These tools are practical applications that can be used in herd MAP control programs.

5.6 References

Blume, J. D. 2002. Likelihood methods for measuring statistical evidence. Stat. Med. 21:2563-2599.

Branscum, A. J., I. A. Gardner, and W. O. Johnson. 2005. Estimation of diagnostic-test sensitivity and specificity through Bayesian modeling. Prev. Vet. Med. 68:145-163.

Chiodini, R. J., H. J. Van Kruiningen, and R. S. Merkal. 1984. Ruminant paratuberculosis (Johne's disease): The current status and future prospects. Cornell Vet. 74:218-262.

Collins, M. 2002. Interpretation of a commercial bovine paratuberculosis enzyme-linked immunosorbent assay by using likelihood ratios. Clin. Diagn. Lab. Immunol. 9:1367-1371.

Collins, M. T., I. A. Gardner, F. B. Garry, A. J. Roussel, and S. J. Wells. 2006. Consensus recommendations on diagnostic testing for the detection of paratuberculosis in cattle in the United States. J. Am. Vet. Med. Assoc. 229:1912-1919.

Collins, M. T., S. J. Wells, K. R. Petrini, J. E. Collins, R. D. Schultz, and R. H. Whitlock. 2005. Evaluation of five antibody detection tests for diagnosis of bovine paratuberculosis. Clin. Diagn. Lab. Immunol. 12:685-692.

Deeks, J. J. and D. G. Altman. 2004. Statistics notes: Diagnostic tests 4: Likelihood ratios. BMJ. 329:168.

Dohoo, I. R., S. W. Martin, and H. Stryhn. 2009. Veterinary Epidemiologic Research. 2nd ed. VER, Inc., Charlottetown, P.E.I., Canada.

Fosgate, G., A. Adesiyun, D. Hird, and S. Hietala. 2006. Likelihood ratio estimation without a gold standard: A case study evaluating a brucellosis c-ELISA in cattle and water buffalo of Trinidad. Prev. Vet. Med. 75:189-205.

Gardner, I. A., S. S. Nielsen, R. J. Whittington, M. T. Collins, D. Bakker, B. Harris, S. Sreevatsan, J. E. Lombard, R. Sweeney, D. R. Smith, J. Gavalchin, and S. Eda. 2011. Consensus-based reporting standards for diagnostic test accuracy studies for paratuberculosis in ruminants. Prev. Vet. Med. 101:18-34.

Garry, F. 2011. Control of paratuberculosis in dairy herds. Vet. Clin. North Am. Food Anim. Pract. 27:599-607.

Georgiadis, M., W. Johnson, I. Gardner, and R. Singh. 2003. Correlation-adjusted estimation of sensitivity and specificity of two diagnostic tests. Appl.Statist. 52, Part 1:63-76.

Hendrick, S. H., D. F. Kelton, K. E. Leslie, K. D. Lissemore, M. Archambault, R. Bagg, P. Dick, and T. F. Duffield. 2006. Efficacy of monensin sodium for the reduction of fecal shedding of Mycobacterium avium subsp. paratuberculosis in infected dairy cattle. Prev. Vet. Med. 75:206-220.

Hendrick, S. H., T. E. Duffield, D. E. Kelton, K. E. Leslie, K. D. Lissemore, and M. Archambault. 2005. Evaluation of enzyme-linked immunosorbent assays performed on milk and serum samples for detection of paratuberculosis in lactating dairy cows. J. Am. Vet. Med. Assoc. 226:424-428.

Hui, S. L. and S. D. Walter. 1980. Estimating the error rates of diagnostic tests. Biometrics. 36:167-171.

Joseph, L., T. W. Gyorkos, and L. Coupal. 1995. Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard. Am. J. Epidemiol. 141:263-272.

Kalis, C. H., M. T. Collins, H. W. Barkema, and J. W. Hesselink. 2004. Certification of herds as free of Mycobacterium paratuberculosis infection: Actual pooled faecal results versus certification model predictions. Prev. Vet. Med. 65:189-204.

Kalis, C. H., J. W. Hesselink, H. W. Barkema, and M. T. Collins. 2000. Culture of strategically pooled bovine fecal samples as a method to screen herds for paratuberculosis. J. Vet. Diagn. Invest. 12:547-551.

Klausen, J., A. Huda, L. Ekeroth, and P. Ahrens. 2003. Evaluation of serum and milk ELISAs for paratuberculosis in Danish dairy cattle. Prev. Vet. Med. 58:171-178.

Lavers, C., H.W. Barkema, I.R. Dohoo, S.L.B. McKenna, and G.P. Keefe. 2013a. Evaluation of milk ELISAs for detection of Mycobacterium avium subspecies paratuberculosis in dairy herds and association with within-herd prevalence. J. Dairy Sci. In press.

Lavers, C., S.L.B. McKenna, I.R. Dohoo, H.W. Barkema, and G.P. Keefe. 2013b. Evaluation of environmental fecal culture for Mycobacterium avium subspecies paratuberculosis detection in dairy herds and association with apparent within-herd prevalence. Can. Vet. J. 54:1053-1060.

Lombard, J. E. 2011. Epidemiology and economics of paratuberculosis. Vet. Clin. North Am. Food Anim. Pract. 27:525-535.

Lombard, J. E., T. M. Byrem, B. A. Wagner, and B. J. McCluskey. 2006. Comparison of milk and serum enzyme-linked immunosorbent assays for diagnosis of Mycobacterium avium subspecies paratuberculosis infection in dairy cattle. J. Vet. Diagn. Invest. 18:448-458.

McKenna, S. L., G. P. Keefe, H. W. Barkema, and D. C. Sockett. 2005. Evaluation of three ELISAs for Mycobacterium avium subsp. paratuberculosis using tissue and fecal culture as comparison standards. Vet. Microbiol. 110:105-111.

Naugle, A. L., W. P. Shulaw, W. J. Saville, T. E. Wittum, and B. Byrum. 2003. Calculation method for likelihood ratios dictates interpretation. Clin. Diagn. Lab. Immunol. 10:729-730.

Nielsen, S. S. 2007. Danish control programme for bovine paratuberculosis. Cattle Pract. 15:161-168.

Nielsen, S. S. 2009. Use of diagnostics for risk-based control of paratuberculosis in dairy herds. In Practice. 31:150-154.

Nielsen, S. S. 2008. Transitions in diagnostic tests used for detection of Mycobacterium avium subsp. paratuberculosis infections in cattle. Vet. Microbiol. 132:274-282.

Nielsen, S. S. and N. Toft. 2008. Ante mortem diagnosis of paratuberculosis: A review of accuracies of ELISA, interferon-gamma assay and faecal culture techniques. Vet. Microbiol. 129:217-235.

Nielsen, S. S. and N. Toft. 2006. Age-specific characteristics of ELISA and fecal culture for purpose-specific testing for paratuberculosis. J. Dairy Sci. 89:569-579.

Nielsen, S. S., C. Gronbaek, J. F. Agger, and H. Houe. 2002. Maximum-likelihood estimation of sensitivity and specificity of ELISAs and faecal culture for diagnosis of paratuberculosis. Prev. Vet. Med. 53:191-204.

Norton, S., W. O. Johnson, G. Jones, and C. Heuer. 2010. Evaluation of diagnostic tests for Johne's disease (Mycobacterium avium subspecies paratuberculosis) in New Zealand dairy cows. J. Vet. Diagn. Invest. 22:341-351.

Sockett, D. C., D. J. Carr, and M. T. Collins. 1992. Evaluation of conventional and radiometric fecal culture and a commercial DNA probe for diagnosis of Mycobacterium paratuberculosis infections in cattle. Can. J. Vet. Res. 56:148-153.

USDA. 2010. Uniform Program Standards for the Voluntary Bovine Johne's Disease Control Program. USDA-APHIS-VS, National Animal Health Monitoring System, Fort Collins, CO (APHIS 91-45-016).

Van Schaik, G., S. Stehman, Y. Schukken, C. Rossiter, and S. Shin. 2003. Pooled fecal culture sampling for Mycobacterium avium subsp. paratuberculosis at different herd sizes and prevalence. J. Vet. Diagn. Invest. 15:233-241.

van Weering, H., G. van Schaik, A. van der Meulen, M. Waal, P. Franken, and K. van Maanen. 2007. Diagnostic performance of the Pourquier ELISA for detection of

antibodies against Mycobacterium avium subspecies paratuberculosis in individual milk and bulk milk samples of dairy herds. Vet. Microbiol. 125:49-58.

Wang, C., B. W. Turnbull, S. S. Nielsen, and Gröhn Y.T. 2011. Bayesian analysis of longitudinal Johne's disease diagnostic data without a gold standard test. J. Dairy Sci. 94:2320-2328.

Wells, S. J., S. M. Godden, C. J. Lindeman, and J. E. Collins. 2003. Evaluation of bacteriologic culture of individual and pooled fecal samples for detection of Mycobacterium paratuberculosis in dairy cattle herds. J. Am. Vet. Med. Assoc. 223:1022-1025.

Whitlock, R. H., S. J. Wells, R. W. Sweeney, and J. Van Tiem. 2000. ELISA and fecal culture for paratuberculosis (Johne's disease): Sensitivity and specificity of each method. Vet. Microbiol. 77:387-398.

Table 5.1. Prior information for milk ELISA and fecal culture test characteristics and study population prevalence used in Bayesian models for estimation of milk ELISA and fecal culture test characteristics for detection of Mycobacterium avium subsp. paratuberculosis infectious (sensitivity) and infected (specificity) cows.

Parameter	Mode (%); 95th percentile (%)	Distribution
Milk ELISA sensitivity	Mode: 30; 95% sure < 85[a]	Beta(1.33, 1.78)
Milk ELISA specificity	Mode: 95; 95% sure > 65[b]	Beta(8.45, 1.39)
Fecal culture sensitivity	Mode: 60; 95% sure < 85[a]	Beta(3.85, 2.90)
Fecal culture specificity	Mode: 98; 95% sure > 70[b]	Beta(9.21, 1.17)
Population prevalence: Infected herds[d]	Mode: 5; 95% sure < 30[c]	Beta(1.51, 10.78)
Population prevalence: Non-infected herds	Prevalence assumed to be 0	Uniform(0,0)

[a]Fecal culture and milk ELISA specificity prior obtained from Norton et al. (2010).
[b]Fecal culture and milk ELISA sensitivity prior obtained from Collins et al. (2006).
[c]Population prevalence priors estimated from Lavers et al. (2013a, 2013b).
[d]Fifteen herds were classified as infected based on at least 1 culture-positive environmental or cow manure sample collected from the herd during the study.

Table 5.2. Cross-tabulated results for two commercial milk enzyme-linked immunosorbent assays[a] (ELISA) and fecal culture (FC) applied to 1829 cows from 15 Mycobacterium avium subsp. paratuberculosis (MAP) infected herds, and 1889 cows from 17 MAP non-infected herds.

	MAP-infected herds				Non-infected herds	
	FC+		FC-		FC-[b]	
	ELISA B+	ELISA B-	ELISA B+	ELISA B-	ELISA B+	ELISA B-
ELISA A+	39	1	7	13	1	9
ELISA A-	6	84	7	1672	5	1874

[a]ELISA A: Parachek2 Mycobacterium paratuberculosis test kit® (Prionics AG, Schlieren-Zurich, Switzerland); ELISA B: Mycobacterium paratuberculosis antibody test kit® (IDEXX, Westbrook, Maine, USA).
[b]By definition, non-infected herds did not have a culture-positive cow result during the study period, therefore all counts for FC+ results in this population were 0.

Table 5.3. Estimation of milk ELISA test characteristics for the detection of Mycobacterium avium subsp. paratuberculosis infectious cows, using a pseudogold standard and a Bayesian latent class model.

Parameter	Pseudogold standard Estimate (95% confidence interval)	Bayesian model Median (95% probability interval)
Milk ELISA A[a]		
Sensitivity[c]	30.8 (23.0-39.5)	28.4 (21.3-36.4)
Specificity[d]	99.5 (99.0-99.7)	99.3 (99.0-99.6)
Milk ELISA B[b]		
Sensitivity	34.6 (26.5-43.5)	33.1 (25.5-41.5)
Specificity	99.7 (99.3-99.9)	99.7 (99.4-99.9)
Fecal culture		
Sensitivity		76.9 (64.4-87.7)
Specificity		99.96 (99.8-1.0)

[a]Parachek2 Mycobacterium paratuberculosis test kit® (Prionics AG, Schlieren-Zurich, Switzerland).
[b]Mycobacterium paratuberculosis antibody test kit® (IDEXX, Westbrook, Maine, USA).
[c]Sensitivity covariance for milk ELISA A and milk ELISA B was 0.17 (0.14-0.20).
[d]Specificity covariance for milk ELISA A and milk ELISA B was 0.0006 (0.00007-0.002).

Table 5.4. Likelihood ratios for diagnosis of Mycobacterium avium subsp. paratuberculosis (MAP) infectious cows, based on five categories of ELISA A[a] corrected optical density (OD) values.

ELISA A OD range[b]	Infectious cows No.	Infectious cows %	Non-infected cows No.	Non-infected cows %	Likelihood ratio (95% Confidence Interval)
< 0.050	88	67.7	1874	99.2	0.68 (0.61-0.77)
0.051-0.099	2	1.5	5	0.3	5.81 (1.14-29.7)
0.100-0.299	9	6.9	7	0.4	18.7 (7.07-49.4)
0.300-0.499	4	3.1	1	0.05	58.1 (6.54-516)
≥ 0.500	27	20.8	2	0.1	196 (47.2-816)
Total	130		1889		

[a]Parachek2 Mycobacterium paratuberculosis test kit® (Prionics AG, Schlieren-Zurich, Switzerland).
[b]Manufacturer cutoff ≥ 0.10.

Table 5.5. Likelihood ratios for diagnosis of Mycobacterium avium subsp. paratuberculosis (MAP) infectious cows, based on five categories of ELISA B[a] sample to positive percentage (S/P %).

	Infectious cows		Non-infected cows		
ELISA B S/P % range	No.	%	No.	%	Likelihood ratio (95% Confidence Interval)
≤ 20	75	57.7	1881	99.6	0.58 (0.50-0.67)
21-39	10	7.7	2	0.1	72.6 (16.1-328)
40-69	7	5.4	3	0.2	33.9 (8.87-130)
70-99	6	4.6	2	0.1	43.6 (8.89-214)
≥ 100	32	24.6	1	0.05	465 (64.0-3380)
Total	130		1889		

[a]Mycobacterium paratuberculosis antibody test kit® (IDEXX, Westbrook, Maine, USA).

Table 5.6. Post-test probabilities (%) for a cow to be Mycobacterium avium subsp. paratuberculosis (MAP) infectious for ELISA A[a] at five ELISA output categories and 6 within-herd MAP prevalence estimates.

		Within-herd true prevalence (Pre-test probability)					
ELISA A OD range	LR	2%	5%	10%	15%	20%	30%
< 0.050	0.7	1.4	3.6	7.2	11.0	14.9	23.1
0.051-0.099	5.8	10.6	23.4	39.2	50.6	59.2	71.3
0.100-0.299	18.7	27.6	49.6	67.5	76.7	82.4	88.9
0.300-0.499	58.1	54.2	75.4	86.6	91.1	93.6	96.1
≥ 0.500	196.2	80.0	91.2	95.6	97.2	98.0	98.8

[a]Parachek2 Mycobacterium paratuberculosis test kit® (Prionics AG, Schlieren-Zurich, Switzerland).

CHAPTER 6

EVALUATION OF REPEATED TEST RESULTS USING A COMMERCIAL MILK ELISA FOR DETECTION OF MYCOBACTERIUM AVIUM SUBSPECIES PARATUBERCULOSIS IN DAIRY CATTLE

6.1 Abstract

Mycobacterium avium ssp. paratuberculosis (MAP) is the causative agent of paratuberculosis, a chronic, infectious enteritis affecting dairy cattle worldwide. In this study, initial milk ELISA tests, followed by repeated tests at 6 or 12 mo intervals, were evaluated for accuracy to detect MAP-infected cows. In addition, factors influencing the probability of detecting a MAP-infected cow, which was initially ELISA negative, on a repeated ELISA were determined. In total, there were 3,145 cows that had ≥ 1 fecal sample and ≥ 1 milk sample collected over 3 rounds of sampling, with each round 6 mo apart. Within this group, 164 (5.2%) cows were fecal culture positive. For a 6 mo test interval, observed Se of the milk ELISA was 22.0% (95% CI: 15.6-28.3%) and 32.6% (95% CI: 26.3-41.2%), for initial and combined initial and repeated tests, respectively. Specificity of an initial ELISA was 99.6% (95% CI: 99.4-99.8%), and 99.2% (95% CI: 98.8-99.5%) for combined tests. For a 12 mo interval, Se of the milk ELISA was 25.6% (95% CI: 18.2-33.0%) and 45.3% (95% CI: 35.6-54.3%) for initial and combined initial and repeated tests, respectively. Specificity of an initial ELISA was 99.6% (95% CI: 99.4-99.9%), and 98.9% (95% CI: 98.4-99.4%) for combined tests. In MAP-infected cows, the magnitude of an initial negative ELISA test was a significant positive predictor for a repeated ELISA to be positive. There was a trend for a MAP-infected cow, initially milk ELISA negative, to be more likely to test ELISA positive with a 12 mo test interval, compared to a 6 mo interval (OR=2.88 (95% CI: 0.84-9.90); P-value 0.09). A repeated milk ELISA test improved the probability to detect a MAP-infected cow, with minimal loss of Sp. A 12 mo test interval provided a greater increase in Se, relative to a single

initial test, than did a 6 mo interval. MAP-infected cows with an initial negative ELISA result close to the cutoff were more likely to be ELISA positive on a repeated milk ELISA test. Cows with repeated positive ELISA results were likely to be MAP-infected and should be a priority for risk management strategies. Given the potential for fluctuation between positive and negative ELISA results, repeated testing within MAP-positive herds improves detection of MAP-infected cows and reduces risk of misclassification based on a single ELISA result.

6.2 Introduction

Paratuberculosis is a chronic, infectious enteritis caused by Mycobacterium avium subspecies paratuberculosis (**MAP**) (Chiodini et al., 1984). There is no perfect ante-mortem test for paratuberculosis (Gardner et al., 2011), creating a challenge for cow-level diagnosis. Diagnostic test accuracy varies with the 3 infection stages: infected (carrying MAP but not infectious); infectious (shedding MAP at the time of testing); and affected (clinical signs present) (Nielsen and Toft, 2008). For infected cattle, sensitivity (**Se**) of the milk ELISA as high as 39% (Nielsen et al., 2002) is reported, with specificity (**Sp**) estimates ranging from 96% (Nielsen et al., 2002) to 99.7% (Collins et al., 2005). For infectious cattle, Se of the milk ELISA ranges from 21% (Lombard et al., 2006) to 61% (Hendrick et al., 2005), and Sp from 95% (Klausen et al., 2003; Hendrick et al., 2005) to 98% (Lombard et al., 2006). Specificity of fecal culture approaches 100% (Nielsen and Toft, 2008), but Se is limited, with estimates ranging from 23% (McKenna et al., 2005) to 29% (Whitlock et al., 2000) for infected cows and 74% for infectious animals (Sockett et

al., 1992). Care should be taken when comparing diagnostic test characteristics, as factors such as target condition, case definition, within-herd prevalence, and statistical methods will all influence estimates (Nielsen and Toft, 2008).

Given the lower cost and faster processing time of the milk ELISA, relative to fecal culture (Tiwari et al., 2006; Collins, 2011), this is the testing option of choice for many veterinarians and dairy producers. Considering the imperfect test characteristics of the milk ELISA, as well as the chronic progressive nature of paratuberculosis, a potential strategy to improve detection of MAP-infected cows is repeated milk ELISA testing. This testing strategy involves repeating a milk ELISA test at a later date and evaluating initial and repeated test results together, rather than attempting to diagnose a cow using a single ELISA result. In cows > 3 yrs of age, repeated milk ELISA results perform significantly better than a single result (Huda et al., 2004). However, it has also been reported that there is variability in repeated serum ELISA tests (Hirst et al., 2002), and that comparing current serum ELISA output to previous results provides minimal advantage over evaluating the current ELISA result (Sweeney et al., 2006).

Repeated milk ELISA testing has also been evaluated using an in-house ELISA (Nielsen, 2002) on samples collected monthly for 3 yrs (Nielsen and Ersboll, 2006). Although monthly testing has the greatest probability to detect a culture positive cow by 4 yrs of age, the study authors point out that a cost-benefit analysis of monthly milk ELISA testing is required. As well, the in-house ELISA in that research is optimized for Se (Nielsen and Ersboll, 2006), and a high number of test-positive milk ELISA results is expected when using this ELISA (Nielsen, 2002; Nielsen and Ersboll, 2006; Nielsen, 2007), making extrapolation of these results to commercial ELISA kits challenging.

Although repeated testing has the potential to increase milk ELISA test Se for detection of MAP-infected cows, before recommending this practice to producers, it is necessary to quantify the information gained from a repeated ELISA test using a commercially available test. It is also necessary to evaluate a testing interval that would potentially be both economically feasible and convenient for producers. A milk ELISA test applied at an annual interval, and at a pre-determined point in a cow's lactation cycle, using a milk sample that is already being collected as part of routine herd milk testing, may be acceptable to many producers from both a convenience and cost perspective.

Factors influencing the outcome of a single milk ELISA result have been investigated. Milk ELISA Se for detection of MAP infection increases with cow age, while Sp decreases (Nielsen and Toft, 2006). As well, milk ELISA test numeric values increase with age (Nielsen and Toft, 2006). The magnitude of an ELISA response also correlates with progression of the disease (Nielsen, 2007). Consideration of factors associated with detection of MAP-infected cows at repeated testing, such as cow age at repeated test, and magnitude of initial ELISA test, may be useful for developing strategies to further improve the probability of detecting MAP-infected cows using a repeated milk ELISA.

The first objective of this study was to investigate the Se and Sp of initial and repeated milk ELISA combinations to detect MAP-infected cows, at both 6 mo and 12 mo test intervals. The second objective was to evaluate factors influencing the probability of a MAP-infected cow to be positive on a repeated ELISA test if it was initially milk-ELISA negative.

6.3 Materials and Methods

6.3.1 Study Design and Sample Collection

Thirty-four herds from Prince Edward Island, New Brunswick and Nova Scotia, Canada, participated in this project. Details on herd selection, herd MAP-status classification, herd demographics, and sample collection have been reported previously (Lavers et al., 2013b). Briefly, herd selection was designed to attain a combination of high-prevalence, low-prevalence, and MAP-negative herds.

The data used in this analysis represent 3 rounds of whole-herd fecal and milk sample collection, gathered at 6 mo intervals, and are referred to as Rounds 1, 2, and 3.(Table 6.1) Individual fecal samples were collected from all lactating cows by study personnel (Lavers et al., 2013b). Individual milk sample collection has been described previously (Lavers et al., 2013a). Briefly, milk samples were also collected 3 times at 6 mo intervals (Rounds 1, 2, and 3) from all lactating cows as part of routine herd milk testing by the regional Dairy Herd Improvement organization (Valacta, Montreal, QC, Canada). Data used in this analysis represent a portion of the data collected in the total project. As described in Lavers et al. (2013b), sample collection included 4 rounds of herd testing. The first round of testing was not performed on all herds, and is not included as part of this analysis. Rounds 1, 2, and 3 in this analysis correspond to the final 3 rounds of testing described in the previous publication (Lavers et al., 2013b). Two of the 34 herds were excluded from the analyses because they did not have 3 herd visits with collection of individual cow fecal or milk samples.

6.3.2 Laboratory Testing: Fecal Culture

Project fecal culture methods have been described previously (Lavers et al., 2013b). Briefly, individual cow fecal samples were pooled by age, with 5 cows in each pool (Kalis et al., 2000). Individual cow samples from positive pools were thawed and cultured individually to determine cow MAP-status. Fecal culture was performed by the technical laboratory staff at the Maritime Quality Milk laboratory (Charlottetown, PEI, Canada), which was USDA-accredited for this technique. Fecal samples were inoculated into ESP para-JEM® broth, according to the manufacturer's protocol, with the exception that samples were incubated for 49 d in the ESP® Culture System II (TREK Diagnostic Systems, Inc., Cleveland, OH, USA), rather than 42 d. Presumptive positive samples (acid-fast positive or positive culture system growth curves) were processed for confirmation with the VetAlert™ Johne's Real-Time PCR kit (Tetracore, Inc., Rockville, MD, USA). A culture sample was considered positive if the presumptive positive sample was confirmed by PCR.

6.3.3 Laboratory Testing: Milk ELISA

Project milk ELISA procedures have been described previously (Lavers et al., 2013a). The indirect ELISA used in this analysis was the Mycobacterium paratuberculosis antibody test kit® (IDEXX, Westbrook, ME, USA). Milk ELISA testing was performed by the technical staff at the Maritime Quality Milk laboratory, which was USDA-accredited for milk ELISA procedures, following the manufacturer's instructions. ELISA output was reported as a sample-to-positive (**S/P**) percentage (%) (((sample optical density (**OD**) – negative control OD)/(positive control OD – negative control OD))*100). Kit instructions describe ELISA output as an S/P ratio. Output described here is the S/P ratio described in the kit, multiplied by a factor of 100. A cow

milk sample was defined as ELISA-positive if S/P% was ≥ 40%, based on the criterion provided in the kit.

6.3.4 Terminology and Statistical Analysis

A herd was considered MAP-positive if ≥ 1 pooled fecal culture collected during any of the 3 herd visits was MAP culture-positive. The target condition for this analysis was a MAP-infected cow, because our primary objective was to determine Se and Sp of repeated milk ELISA testing for MAP-infected cows, not MAP-infectious or MAP-affected cows. A cow was considered MAP-infected if 1 or more fecal samples collected from her during the project (minimum 1 and maximum 3 fecal samples collected per cow) was MAP culture-positive. Requiring fecal shedding and antibody response to occur at the same point in time would indicate MAP-infectiousness for the ELISA, which was not our target condition for this study.

Within MAP-infected cows, the Se of an initial milk ELISA test was defined as the probability of the initial ELISA being positive, given the cow was MAP-infected (e.g. 0.20, therefore the probability of an initial test being negative would be (1 – 0.20) = 0.80). The combined initial and repeated milk ELISA tests were evaluated in parallel, where the combined test result was considered positive if either the initial or repeated test, or both tests, was positive (Dohoo et al., 2009). The Se of combined tests was calculated as the remainder of the product of two probabilities (see Equation 1): (1) probability of an initial test being negative (e.g. 0.80); and (2) probability of a repeated test being negative, given the initial test was negative and the repeated test was not missing (e.g. 0.70).

$$Se = 1 - p(T_1-) * p(T_2- | T_1-) \qquad \text{Equation 1}$$

Using the example probabilities, the combined Se would be 1 – [(0.80)*(0.70)] = 0.44. Initial and combined sensitivities were calculated separately for data collected from 6 and 12 mo test intervals. In order to optimize the data utilization, if the first test (Round 1) was missing, the subsequent sample (Round 2) was used for initial Se, creating a larger dataset for 6 mo than 12 mo interval Se and Sp data (Testing scenarios 2 and 6, Table 6.1).

Within MAP-negative cows, the Sp of an initial milk ELISA test was defined as the probability of the initial ELISA test being negative, given the cow was MAP-negative (e.g. 0.98). The Sp of combined initial and repeated milk ELISA tests was calculated as the product of two probabilities (see Equation 2): (1) probability of initial test being negative (e.g. 0.98); and (2) probability of repeated test being negative, given the initial test was negative and the repeated test was not missing (e.g. 0.99).

$$Sp = p(T_1-) * p(T_2-|T_1-) \qquad \text{Equation 2}$$

Using the example probabilities, the combined Sp would be (0.98)*(0.99) = 0.97. Again, initial and combined specificities were calculated separately for data collected from 6 and 12 mo test intervals.

Confidence intervals (95%) for test characteristics of an initial test were calculated using the large-sample approximation. Confidence intervals for 2 test combinations (product of a proportion) were computed using 2.5% and 97.5% percentile values from parametric bootstrapped estimates based on 1000 resamples.

A logistic regression analysis determined predictors impacting the probability that a MAP-infected cow, initially milk ELISA-negative, tested positive on the repeated milk ELISA. Predictors evaluated were: INTERVAL, a dichotomous variable indicating a 6

mo (0) or 12 mo (1) test interval; AGE, the cow age in days at initial test and centered with 2 yrs as the baseline; S/P, S/P % of the initial milk ELISA; and WHP, the mean within-herd MAP prevalence, based on fecal culture. Initially, predictors were tested within a logistic mixed model with random effects at the herd and cow levels, accounting for the 3-level hierarchical structure of the data (herd/cow/repeated tests). However, clustering at the herd-level and cow-level was minimal (see results), and only cow-level factors were significant, therefore, a generalized estimating equation (**GEE**) model was selected as the best option for handling the hierarchical data. This GEE model accounted for repeated milk ELISA sampling at the cow-level, which allowed for accurate estimation of 95% confidence intervals around estimates, and produced population averaged estimates. Univariable analyses were initially performed, and subsequently, all factors were evaluated within a multivariable model using backward selection. Regardless of the P-value, AGE was forced into the model because it was a potential confounder for the effect of INTERVAL on the outcome. INTERVAL was a key variable of interest, and was therefore included in the model, regardless of P-value. Linear relationships between continuous predictors and the outcome were determined using fractional polynomial models, in order to determine if transformations (e.g. quadratic) of the predictors provided a better fit for the model. Correlation between repeated observations within a cow was calculated. The analyses were conducted using Stata/IC® Version 11.2 (StataCorp LP, College Station, TX, USA). A P-value ≤ 0.05 was considered significant for inclusion in the final model, with the exception of variables that were forced into the model (above).

6.3.5 Data Management

For a cow's data to be included in this dataset, the cow had to have ≥ 1 fecal culture result and ≥ 1 milk ELISA result. For the 32 herds with 3 rounds of testing, there were 4,145 cows that had ≥ 1 fecal culture or ELISA result throughout the 3 rounds of testing. There were 288 cows with no fecal culture data and 269 cows with no ELISA data which were not included in the dataset. If a cow's first ELISA result came from the final round of testing (Round 3) she was excluded from the dataset because she did not have the opportunity to have a repeated ELISA (Testing scenario 7, Table 6.1). There were 443 cows with a single ELISA test in Round 3 (5 MAP-infected and 438 MAP-negative cows). Applying these criteria, the final dataset for analysis included 3,145 cows.

For test characteristics of initial and repeated tests, the chronologically first milk ELISA test for a cow was considered the initial test, and the result from the next round of testing was the repeated test. If the cow was not repeat tested, the repeated result was 'not tested'.

The 6 mo test interval data came from a combination of Rounds 1 (initial) and 2 (repeated) and Rounds 2 (initial) and 3 (repeated) combinations. If a cow had data from both combinations (i.e. 3 ELISA results available), the first chronological pair was included in the estimation of Se and Sp (Testing scenario 1, Table 6.1). This was to allow for calculation of confidence intervals, which was not possible if all data were included due to dependence of test results within a cow. The 6 mo repeated ELISA had to be ≥ 110 d and ≤ 255 d following the initial ELISA. For analysis of a 12 mo test interval, the only possible test combination was Rounds 1 (initial) and 3 (repeated) (Table 6.1). The repeated ELISA had to be ≥ 290 d and ≤ 430 d following the initial ELISA.

For the GEE models, repeated tests within a cow could be accounted for, and therefore all test combinations were included. All complete test combinations were eligible for inclusion in the model, provided the cow was MAP-infected and the initial milk ELISA result was negative. In a subset of cows that had data for all 3 rounds, data from Rounds 1 and 2 (6 mo interval) and data from Rounds 1 and 3 (12 mo interval) were included.

6.4 Results

6.4.1 Descriptive Statistics

Median herd size was 66 milking cows (mean: 82; range: 28 to 220). The median cow age at testing was 4.0 y (mean: 4.4 y; range: 1.8 – 17.3 y). In total, there were 3,145 cows that were included in the final dataset for analyses. Within this group, 164 (5.2%) cows were MAP-infected. Within the MAP-infected cows, 84% had ≥ 2 fecal samples collected, and 73% had ≥ 2 milk samples collected. Within the MAP-negative cows, 80% had ≥ 2 fecal samples collected, and 82% had ≥ 2 milk samples collected (Table 6.2).

There were 26 (16%) MAP-infected cows with 1 fecal culture result. There were 76 (46%) MAP-infected cows with 2 fecal culture results; 51 were culture-positive on 1 of 2 fecal samples, and 25 were culture-positive on both fecal samples. There were 62 (38%) MAP-infected cows with 3 fecal culture results; 35 were culture-positive on 1 of 3 fecal samples, 16 were culture-positive on 2 of 3 fecal samples, and 11 were culture-positive on all 3 fecal samples.

In the complete dataset, there were 21 test pairs where both initial and repeated milk ELISA tests were positive. These test pairs came from 16 cows, accounting for cows

having 3 tests and being positive on every test. Fifteen of these 16 cows were MAP-infected. The mean S/P% of the initial ELISA in these 15 cows was 136% (range: 44-252%). Seventy-five percent of these 15 MAP-infected cows had S/P% values ≥ 70% on the initial ELISA.

6.4.2 Test Characteristics of Initial and Repeated Milk ELISA Tests

Initial and repeated milk ELISA results from a 6 mo test interval are displayed in Table 6.3. Mean interval between tests was 174 d. Observed Se of an initial ELISA test was 22.0% (95% CI: 15.6-28.3%). Observed Se of combined initial and repeated ELISA tests was 32.6% (95% CI: 26.3-41.2%). Observed Sp of an initial ELISA test was 99.6% (95% CI: 99.4-99.8%), and observed Sp of combined ELISA tests in a 6 mo interval was 99.2% (95% CI: 98.8-99.5%).

Initial and repeated milk ELISA results from a 12 mo test interval are displayed in Table 6.4. Mean interval between tests was 364 d. Observed Se of an initial ELISA test was 25.6% (95% CI: 18.2-33.0%). Observed Se of combined initial and repeated ELISA tests was 45.3% (95% CI: 35.6-54.3%). Observed Sp of an initial ELISA test was 99.6% (95% CI: 99.4-99.9%), and observed Sp of combined ELISA tests in a 12 mo interval was 98.9% (95% CI: 98.4-99.4%). The initial Se and Sp estimates for 6 mo and 12 mo intervals were calculated using the same methods, but because the 12 mo dataset was slightly smaller, separate initial Se and Sp were calculated from the dataset for each interval, in order to provide comparisons for the combined results.

6.4.3 Predictors Associated with Repeated Positive Milk ELISA Results in MAP-Infected Cows

There were 108 MAP-infected cows with an initial negative milk ELISA result and at least 1 repeated ELISA test. Sixty-three cows contributed 1 pair of ELISA tests (initial and repeated test), and 45 cows contributed 2 pairs. As noted in Materials and Methods, a logistic mixed model with random effects demonstrated little clustering at the herd-level (variance: $2.84*10^{-8}$), or at the cow-level (variance: $9.89*10^{-6}$) (data available upon request). Results from the GEE model indicated that the S/P% value of an initial negative ELISA result was a significant predictor for a repeated positive ELISA result in MAP-infected cows (Table 6.5). For every 5% increase in initial S/P% value of a negative test, a MAP-infected cow was 2.3 times more likely to test positive on a repeated ELISA test. There was a trend for a MAP-infected cow to be more likely positive on a repeated ELISA test with a 12 mo test interval, compared to a 6 mo interval (OR=2.88; P-value 0.09).

6.5 Discussion

Combined information from initial and repeated milk ELISA tests increased the probability of detecting a MAP-infected cow, compared to an initial ELISA test only. For 6 mo, and in particular 12 mo test intervals, a repeated ELISA provided considerable gain in Se with a relatively small decrease in Sp. When a 6 mo test interval was utilized, the Se of combined initial and repeated tests was 32.6%, a 48% increase over the initial test Se. Using a 12 mo interval, the probability to detect a MAP-infected cow was 45.3%, a 77% increase over the initial test Se. Specificity of combined tests for a 6 mo interval was only half a percentage point lower than initial test Sp, and 0.7 of a percentage point lower

using a 12 mo interval. Huda et al. (2004) also reported that repeated milk ELISA testing increased the probability of detecting a MAP-infected cow. Using an in-house ELISA with various cutoffs dependent on cow age, Huda et al. (2004) found that area under the curve was significantly greater for receiver operating characteristic curves representing 3 repeated milk ELISA samples, compared to 1 sample, for cows ≥ 3 yrs of age. However, the use of an in-house ELISA precluded extrapolation of results to commercial milk ELISA tests.

Further investigation of test intervals, utilizing a GEE model, demonstrated a trend for increased probability to detect a MAP-infected cow using a 12 mo test interval, as compared to a 6 mo interval (OR=2.7; P=0.11). Impact of an increased test interval was also suggested in a serum ELISA study, which noted a longer interval could result in more cows seroconverting from negative to positive ELISA status (Hirst et al., 2002). Small sample size in the current study may have contributed to lack of statistical significance, and a larger sample size may allow determination as to whether the difference between test intervals is significant at P=0.05. Paratuberculosis is a chronic, progressive disease, and it is expected that the probability for a test to detect disease will increase as an animal ages (Nielsen and Ersboll, 2006). Milk ELISA detection of MAP-infection increases approximately linearly from 2 to 5 yrs of age (Nielsen and Toft, 2006). In the current dataset, age was not a significant predictor for a MAP-infected cow, initially ELISA negative, to be positive on a repeated ELISA. Lack of significance may have been related to the relatively small number of MAP-infected cows in the dataset.

The impact of a repeated test on Sp in the current study was minimal, however the number of repeated tests on the same animals was minimal. Nielsen and Ersboll (2006)

used data collected from monthly milk ELISA tests performed over a 3 yr period. With monthly testing, by 4 yrs of age, a cow that was never culture-positive had an 18% probability of having a positive milk ELISA test (approximated from a figure in that paper). A lower Sp due to high test frequency may be acceptable when producers and veterinarians anticipate and understand this likelihood. For example, in the Danish Control Programme for Bovine Paratuberculosis (Nielsen, 2007), where cattle are milk ELISA tested 4 times per year, communication and education is a key part of program success. However, if a high test frequency is applied without the understanding that overall Sp will be lowered, test results will be frustrating for the producer to interpret.

A 12 mo test interval might be more financially acceptable and convenient to herds undertaking a MAP control program. Nielsen and Ersboll (2006) concluded that although monthly testing in cows < 4 yrs of age would increase the Se of the ELISA, a cost-benefit analysis to determine optimal test frequency is required (Nielsen and Ersboll, 2006). An annual test would fit conveniently within current production systems, as many preventative health measures (e.g. vaccinations and dry cow therapy) are based on the lactation cycle of the cow.

The quantitative value (S/P %) of an initial negative milk ELISA test was predictive of a MAP-infected cow being detected on a repeated milk ELISA test. This is in agreement with an earlier study evaluating single serum ELISA results on repository samples, which found that cows with OD values just below the cutoff (0.050 - < 0.100) were 15 times more likely to be MAP-infected than non-infected (Collins and Sockett, 1993). Regarding ELISA values above the cutoff, the magnitude of the quantitative value may correlate with the infectious status (Nielsen, 2007), and it has been reported that

moderate and heavy MAP shedders have significantly higher quantitative ELISA values than culture-negative cows (van Schaik et al., 2003). The relationship between quantitative values and milk production has also been studied, and a negative relationship between milk production and OD of the ELISA test was reported (Kudahl et al., 2004). Together, these results suggest it is beneficial to consider the magnitude of the ELISA value, in addition to the dichotomous result. If, for example, a producer is unable to repeated testing on all ELISA negative cows, the quantitative value may be useful to target repeated testing to cows with a value close to the cutoff. Collins and Sockett (1993) also suggest that quantitative values may be useful when prioritizing culling decisions.

Within this dataset, there were 16 cows with a positive milk ELISA result on both initial and repeated tests. Fifteen of the 16 cows (94%) with a positive-positive pattern were MAP-infected. The 1 MAP-negative cow with a positive-positive pattern originated from a MAP-infected herd that had a mean fecal culture within-herd prevalence of almost 10%. It is possible that this cow was truly a MAP-infected cow with false-negative fecal culture results. These results suggest that repeated positive ELISA tests represent a high likelihood for a cow to be MAP-infected. Other publications are also supportive of this likelihood, and it has been noted that cows with repeated positive ELISA results are more likely to be shedding MAP in the near future than cows with fluctuating antibody response profiles (Nielsen, 2008). One of the recommendations in Denmark's MAP control program is to cull cows with 2 positive ELISA results when testing 4 times per year (Nielsen, 2009). It has also been reported that culling repeatedly ELISA positive cows was 1 of 3 management strategies that was found to significantly decrease ELISA within-herd prevalence (Nielsen and Toft, 2011). In addition to repeated positive results

representing a high likelihood of MAP-positivity, it has been noted that cows with repeated positive results, as well as cows with their last test positive and previous tests negative for MAP, produced significantly less milk than cows that were repeatedly milk ELISA negative (Nielsen et al., 2006). When interpreting repeated milk ELISA test results, cows with repeated positive results are more likely to be infected with MAP and therefore should be managed as MAP-infected within a herd risk management program.

Fluctuation of serum ELISA results from positive to negative has been described (Hirst et al., 2002; Sweeney et al., 2006), and is a potential source of confusion for producers and veterinarians. Amongst 21 cows with an initial positive milk ELISA result, 47% (10/21) had a negative ELISA 6 mo later. The rate of reversion in MAP-negative cows was 100% (9/9), while only 1 of 12 of MAP-infected cows reverted to a negative ELISA status. Reversion proportions for the 12 mo interval were almost identical. These results are similar to serum ELISA studies. Hirst et al. (2002) reported that 40% (62/157) of seropositive cows were negative on a repeated ELISA test. However, the fecal culture status of these animals was not known. In another longitudinal serum ELISA study, where fecal culture results were available, Sweeney et al. (2006) found that although 95% (17/18) of culture-negative cows with an initial positive serum result had a repeated negative serum result, only 24% (14/58) of culture-positive cows reverted from seropositive to seronegative. Reversion from positive to negative on a repeated ELISA appears to occur more commonly in MAP-negative cows. However, as recommended in the Danish paratuberculosis control program, these cows should be regarded as low risk, and further repeated testing is recommended to better establish their MAP-infection status (Nielsen, 2007).

The observational structure of this project lends both strengths and weaknesses. A strength was that cows studied were naturally infected within a selection of MAP-negative, low-prevalence and high-prevalence herds. The study cow population would be representative of many herds in the North American industry, where most MAP-positive herds are expected to be low to moderate prevalence (Lombard, 2011). Therefore selection bias at the cow-level in this population should have had little impact on our results. The primary weaknesses of the observational study were that cows provided variable numbers of samples, and the large sample size precluded tissue culture at postmortem, which would have allowed for further determination of true MAP status. As well, despite the large sample size, the subset of MAP-infected cows was small because MAP is a disease of low prevalence in most herds.

6.5.1 Conclusions

A repeated milk ELISA test substantially improved the probability to detect a MAP-infected cow, with minimal loss of Sp. The data indicated a trend for a 12 mo test interval provided a greater increase in Se, relative to a single initial test, than did a 6 mo interval, although this finding was not statistically significant. MAP-infected cows with a negative ELISA result close to the cutoff were more likely to be ELISA positive on a repeated milk ELISA test. Cows with repeated positive ELISA results were likely to be MAP-infected and should be a priority for risk management strategies. Given the potential for fluctuation between positive and negative ELISA test results, repeated testing within a MAP-positive herd provides an improved understanding of a cow's MAP status, and reduces risk of misclassification based on a single test result.

6.6 References

Chiodini, R. J., H. J. Van Kruiningen, and R. S. Merkal. 1984. Ruminant paratuberculosis (Johne's disease): The current status and future prospects. Cornell Vet. 74:218-262.

Collins, M. T. 2011. Diagnosis of paratuberculosis. Vet. Clin. North Am. Food Anim. Pract. 27:581-91.

Collins, M. T. and D. C. Sockett. 1993. Accuracy and economics of the USDA-licensed enzyme-linked immunosorbent assay for bovine paratuberculosis. J. Am. Vet. Med. Assoc. 203:1456-1463.

Collins, M. T., S. J. Wells, K. R. Petrini, J. E. Collins, R. D. Schultz, and R. H. Whitlock. 2005. Evaluation of five antibody detection tests for diagnosis of bovine paratuberculosis. Clin. Diagn. Lab. Immunol. 12:685-692.

Gardner, I. A., S. S. Nielsen, R. J. Whittington, M. T. Collins, D. Bakker, B. Harris, S. Sreevatsan, J. E. Lombard, R. Sweeney, D. R. Smith, J. Gavalchin, and S. Eda. 2011. Consensus-based reporting standards for diagnostic test accuracy studies for paratuberculosis in ruminants. Prev. Vet. Med. 101:18-34.

Hendrick, S. H., T. E. Duffield, D. E. Kelton, K. E. Leslie, K. D. Lissemore, and M. Archambault. 2005. Evaluation of enzyme-linked immunosorbent assays performed on milk and serum samples for detection of paratuberculosis in lactating dairy cows. J. Am. Vet. Med. Assoc. 226:424-428.

Hirst, H. L., F. B. Garry, and M. D. Salman. 2002. Assessment of test results when using a commercial enzyme-linked immunosorbent assay for diagnosis of paratuberculosis in repeated samples collected from adult dairy cattle. J. Am. Vet. Med. Assoc. 220:1685-1689.

Huda, A., G. Jungersen, and P. Lind. 2004. Longitudinal study of interferon-gamma, serum antibody and milk antibody responses in cattle infected with Mycobacterium avium subsp. paratuberculosis. Vet. Microbiol. 104:43-53.

Kalis, C. H., J. W. Hesselink, H. W. Barkema, and M. T. Collins. 2000. Culture of strategically pooled bovine fecal samples as a method to screen herds for paratuberculosis. J. Vet. Diagn. Invest. 12:547-551.

Klausen, J., A. Huda, L. Ekeroth, and P. Ahrens. 2003. Evaluation of serum and milk ELISAs for paratuberculosis in Danish dairy cattle. Prev. Vet. Med. 58:171-178.

Kudahl, A., S. S. Nielsen, and J. T. Sorensen. 2004. Relationship between antibodies against Mycobacterium avium subsp. paratuberculosis in milk and shape of lactation curves. Prev. Vet. Med. 62:119-134.

Lavers, C., H.W. Barkema, I.R. Dohoo, S.L.B. McKenna, and G.P. Keefe. 2013a. Evaluation of milk ELISAs for detection of Mycobacterium avium subspecies paratuberculosis in dairy herds and association with within-herd prevalence. J. Dairy Sci. In press.

Lavers, C., S.L.B. McKenna, I.R. Dohoo, H.W. Barkema, and G.P. Keefe. 2013b. Evaluation of environmental fecal culture for Mycobacterium avium subspecies paratuberculosis detection in dairy herds and association with apparent within-herd prevalence. Can. Vet. J. 54:1053-1060.

Lombard, J. E. 2011. Epidemiology and economics of paratuberculosis. Vet. Clin. North Am. Food Anim. Pract. 27:525-535.

Lombard, J. E., T. M. Byrem, B. A. Wagner, and B. J. McCluskey. 2006. Comparison of milk and serum enzyme-linked immunosorbent assays for diagnosis of Mycobacterium avium subspecies paratuberculosis infection in dairy cattle. J. Vet. Diagn. Invest. 18:448-458.

McKenna, S. L., G. P. Keefe, H. W. Barkema, and D. C. Sockett. 2005. Evaluation of three ELISAs for Mycobacterium avium subsp. paratuberculosis using tissue and fecal culture as comparison standards. Vet. Microbiol. 110:105-111.

Nielsen, S. S. 2009. Use of diagnostics for risk-based control of paratuberculosis in dairy herds. In Pract. 31:150-154.

Nielsen, S. S. 2008. Transitions in diagnostic tests used for detection of Mycobacterium avium subsp. paratuberculosis infections in cattle. Vet. Microbiol. 132:274-282.

Nielsen, S. S. 2007. Danish control programme for bovine paratuberculosis. Cattle Pract. 15:161-168.

Nielsen, S. S. 2002. Variance components of an enzyme-linked immunosorbent assay for detection of IgG antibodies in milk samples to Mycobacterium avium subspecies paratuberculosis in dairy cattle. J. Vet. Med. 49:384-387.

Nielsen, S. S. and N. Toft. 2011. Effect of management practices on paratuberculosis prevalence in Danish dairy herds. J. Dairy Sci. 94:1849-1857.

Nielsen, S. S. and N. Toft. 2008. Ante mortem diagnosis of paratuberculosis: A review of accuracies of ELISA, interferon-gamma assay and faecal culture techniques. Vet. Microbiol. 129:217-235.

Nielsen, S. S. and A. K. Ersboll. 2006. Age at occurrence of Mycobacterium avium subspecies paratuberculosis in naturally infected dairy cows. J. Dairy Sci. 89:4557-4566.

Nielsen, S. S. and N. Toft. 2006. Age-specific characteristics of ELISA and fecal culture for purpose-specific testing for paratuberculosis. J. Dairy Sci. 89:569-579.

Nielsen, S., C. Enevoldsen, and N. Toft. 2006. Milk production losses associated with bovine paratuberculosis diagnosed from repeated testing. Pages 6-11 in Proceedings of the 11th International Symposium on Veterinary Epidemiology and Economics, Cairns, Australia.

Nielsen, S. S., C. Gronbaek, J. F. Agger, and H. Houe. 2002. Maximum-likelihood estimation of sensitivity and specificity of ELISAs and faecal culture for diagnosis of paratuberculosis. Prev. Vet. Med. 53:191-204.

Sockett, D. C., D. J. Carr, and M. T. Collins. 1992. Evaluation of conventional and radiometric fecal culture and a commercial DNA probe for diagnosis of Mycobacterium paratuberculosis infections in cattle. Can. J. Vet. Res. 56:148-153.

Sweeney, R. W., R. H. Whitlock, S. McAdams, and T. Fyock. 2006. Longitudinal study of ELISA seroreactivity to Mycobacterium avium subsp. paratuberculosis in infected cattle and culture-negative herd mates. J. Vet. Diagn. Invest. 18:2-6.

Tiwari, A., J. A. VanLeeuwen, S. L. McKenna, G. P. Keefe, and H. W. Barkema. 2006. Johne's disease in Canada part I: Clinical symptoms, pathophysiology, diagnosis, and prevalence in dairy herds. Can. Vet. J. 47:874-882.

van Schaik, G., C. R. Rossiter, S. M. Stehman, S. J. Shin, and Y. H. Schukken. 2003. Longitudinal study to investigate variation in results of repeated ELISA and culture of fecal samples for Mycobacterium avium subsp paratuberculosis in commercial dairy herds. Am. J. Vet. Res. 64:479-484.

Whitlock, R. H., S. J. Wells, R. W. Sweeney, and J. Van Tiem. 2000. ELISA and fecal culture for paratuberculosis (Johne's disease): Sensitivity and specificity of each method. Vet. Microbiol. 77:387-398.

Table 6.1. Representation of the origin of data used in the calculation of sensitivity and specificity of initial and combined initial and repeated (6 month and 12 month interval) milk ELISA tests for the detection of Mycobacterium avium subsp. paratuberculosis. Milk samples were collected every 6 months, for a total of 3 rounds of testing, from 32 herds in the Maritime provinces.

				6 Month Interval		12 Month Interval	
	Round of Testing			Initial	Combined	Initial	Combined
Cow Testing Scenario	1	2	3	Round(s) of Testing Utilized			
1	Tested	Tested	Tested	1	1 and 2	1	1 and 3
2	NT[a]	Tested	Tested	2	2 and 3	n/a[b]	n/a
3	Tested	NT	Tested	1	n/a	1	1 and 3
4	Tested	Tested	NT	1	1 and 2	1	n/a
5	Tested	NT	NT	1	n/a	1	n/a
6	NT	Tested	NT	2	n/a	n/a[b]	n/a
7	NT	NT	Tested	n/a	n/a	n/a[b]	n/a

[a] NT = Cow was not tested.
[b] n/a = Not applicable. If an initial test was not available, the initial and combined test characteristics were not calculated. If a repeat test was not available, the combined test characteristics were not calculated.

Table 6.2. Frequency of Mycobacterium avium subsp. paratuberculosis fecal culture and milk ELISA test events for 3,145 cows from 32 Atlantic Canadian dairy herds.

		Fecal Culture Tests					
		MAP-infected (164 cows)			MAP-negative (2,981 cows)		
		1	2	3	1	2	3
Milk ELISA Tests	1	16	24	3	304	164	71
	2	9	32	30	258	622	501
	3	1	20	29	25	415	621
	Total	26	76	62	587	1201	1193

Table 6.3. Initial and repeated Mycobacterium avium subsp. paratuberculosis milk ELISA test results collected at a 6 month test interval, for 164 MAP-infected and 2981 MAP-negative cows, with a minimum of 140 days and maximum of 253 days between initial and repeated tests.

Milk ELISA Result		Fecal Culture Classification	
Initial Test	Repeated Test	MAP-infected	MAP-negative
Negative	Negative	74	2068
Negative	Positive	14	9
Negative	Not Tested	40	892
Positive	Positive	11	0
Positive	Negative	1	9
Positive	Not Tested	24	3
	Total	164	2981

Table 6.4. Initial and repeated Mycobacterium avium subsp. paratuberculosis milk ELISA test results collected at a 12 month test interval for 133 MAP-infected and 2183 MAP-negative cows, with a minimum of 293 days and maximum of 426 days between initial and repeated tests.

Milk ELISA Result		Fecal Culture Classification	
Initial Test	Repeated Test	MAP-infected	MAP-negative
Negative	Negative	47	1321
Negative	Positive	17	10
Negative	Not Tested	35	846
Positive	Positive	6	0
Positive	Negative	1	5
Positive	Not Tested	27	3
	Total	133	2185

Table 6.5. Odds ratios for generalized estimating equation model predictors influencing the probability for a Mycobacterium avium subsp. paratuberculosis culture positive cow with an initial negative milk ELISA test to have a repeated positive milk ELISA test.

Predictor	Coefficient	Odds Ratio (OR)	OR 95% Confidence Interval	P-value
Intercept[a]	-2.86		-4.28- -1.47 (95% CI for coefficient)	
Test interval (0=6 mo; 1=12 mo)	1.06	2.88	0.84-9.90	0.09
Cow age[b] at initial ELISA (days)	0.0003	1.0003	0.9995-1.001	0.40
Initial ELISA S/P %	0.16	1.18	1.08-1.28	<0.001

[a] Correlation between repeated observations within a cow = -0.03.
[b] Cow age centered with 2 yr as baseline.

CHAPTER 7. SUMMARIZING CONCLUSIONS

Johne's disease, or paratuberculosis, has long been recognized by the veterinary community as a serious disease affecting cattle (Chiodini et al., 1984). Despite a disease history spanning almost 200 years, the causative bacterium, Mycobacterium avium subspecies paratuberculosis (**MAP**), continues to challenge researchers in regards to diagnosis, treatment, and control. There is no effective treatment or vaccination available to cure or prevent paratuberculosis infection, and therefore management of the disease relies on successful MAP control programs (Garry, 2011). The United States' National Voluntary Bovine Johne's Disease Control Program suggests 3 sequential steps for a MAP control program: producer education, followed by development of a risk assessment and management plan, and finally establishment of a cow-level testing program in herds established as MAP-positive (Whitlock, 2010). When cow-level test results are used to implement management changes for MAP-infectious cows, the return on investment generally justifies testing costs (Dorshorst and Lombard, 2006).

The focus of the research reported in this thesis was the evaluation of diagnostic tests for both herd-level (environmental culture and milk ELISA) and cow-level (milk ELISA) detection of paratuberculosis. One of the basic tenets of this research project was the desire to evaluate these tests within a relatively large population of herds that represented a spectrum of MAP-negative, low prevalence and moderate to high prevalence herds. Although paratuberculosis is a disease with a low to moderate within-herd prevalence (Lombard, 2011), many paratuberculosis diagnostic test evaluations have

targeted moderate and high prevalence herds to maximize the efficient use of limited research dollars. However, the disadvantage of this strategy is that study populations are often not reflective of target populations in regard to MAP within-herd prevalence. The 32 herds that contributed data to this analysis were therefore purposively selected, in an effort to obtain a sound mixture of MAP-negative, low prevalence and high prevalence herds.

7.1 Establishment of a Case Definition

Herds were classified as MAP-positive and MAP-negative based on repeated pooled fecal cultures, performed on samples collected from all lactating cows in a herd. Repeated sampling, in the form of multiple negative tests from all adult cattle in the herd, is required for an accurate establishment of a MAP-negative herd status (Collins et al., 2005). All 18 study herds classified as MAP-negative had 4 whole-herd pooled fecal culture collections with no positive pooled fecal culture results. As well, although they were not designated as closed herds, none of the MAP-negative herds increased in size during the study, making it unlikely that they introduced new cows into the milking herd that were MAP-positive. It was assumed in this analysis that MAP herd-status was stable over the study period. There is limited field data supporting the short-term efficacy of MAP eradication programs, even in MAP-positive low-prevalence herds (Collins et al., 2006). As a result, a change from MAP-positive to MAP-negative herd status was also considered unlikely in the 2 yr study time frame.

Repeated sampling will also maximize the identification of low prevalence MAP-positive herds (Kalis et al., 2004) and thereby minimize the influence of misclassification bias on herd sensitivity and specificity (Christensen and Gardner, 2000). In our dataset,

false-negative herd classification, based on cross-sectional sampling at 1 point in time rather than evaluation of repeated herd tests, was most common in low-prevalence study herds. The use of repeated sampling to establish the herd case definition, therefore, maximized the correct classification of both MAP-positive and MAP-negative herds.

7.2 Herd-level Detection of MAP

7.2.1 Herd-level Detection of MAP: Environmental Culture

The use of environmental culture to detect MAP-positive herds was investigated in Chapter 2. Using generalized estimating equation (**GEE**) logistic models, both overall herd sensitivity and herd specificity were estimated, and factors influencing these herd-level test characteristics were also investigated. The only significant relationship was a positive association between fecal culture within-herd prevalence and sensitivity of environmental culture.

Environmental culture sensitivity is expected to be higher in high-prevalence herds (Smith et al., 2011). This was consistent with our study results, and although environmental culture sensitivity was only 51% when within-herd MAP prevalence was 2%, it increased quickly, and approached 100% sensitivity at moderate within-herd prevalence levels of approximately 8%.

Repeated herd testing is frequently used in control programs to determine MAP herd-status (Whitlock, 2010). Although not statistically significant, the greatest increase in sensitivity of repeated environmental culture sets occurred when 2 sets of environmental culture samples, collected 3 mo apart, were evaluated in parallel, as compared to a single environmental culture set. Given the lower sensitivity of

environmental culture in low-prevalence herds, it is advisable to repeat environmental cultures to ensure the appropriate classification of MAP-positive low-prevalence and MAP-negative herds.

The proportion of positive environmental culture samples within an environmental culture set was positively associated with the herd's apparent fecal culture within-herd prevalence, extending the applicability of this herd-level test beyond a dichotomous herd MAP diagnosis. It is advantageous if a herd-level paratuberculosis test can give a crude estimation of within-herd MAP prevalence, as this information would help guide the development of a MAP control program (Garry, 2011). For example, aggressive testing may be more beneficial and cost-effective in high-prevalence herds (Wells et al., 2003).

The sampling protocol tested in Chapter 2 was modeled after the protocol designed by the Voluntary Bovine Johne's Disease Control Program (USDA, 2010). This protocol was intended to be a realistic sampling procedure, in order that estimates produced in the analysis would reflect what could be expected in the practical application of MAP control programs. When evaluating collection location in all our study herds, manure storage and concentration areas had a numerically greater number of positive environmental culture samples than cow concentration areas, although this was not statistically significant. Previous studies have also reported manure storage and shared alleyways as the sites most likely to be positive (Raizman et al., 2004; Pillars et al., 2009; Smith et al., 2011). Collecting environmental samples from manure storage and concentration sites will provide the optimum chance to detect MAP.

Environmental culture is an acceptable herd test for classification of MAP herd-status in MAP-negative and MAP-positive herds. Herd sensitivity of environmental culture is significantly impacted by the within-herd prevalence of MAP. The proportion of positive environmental culture samples within a sample set provides a crude estimation of within-herd prevalence which is useful information in the development of MAP control programs.

7.2.2 Herd-level Detection of MAP: Milk ELISA

Assessment of cow-level milk ELISA results, compiled to form a herd-level MAP diagnosis, was completed in Chapter 3. Using null GEE logistic models, both overall herd sensitivity and specificity were estimated. In a multivariable GEE model, factors influencing herd-level test characteristics were investigated. The only significant association was a positive association between fecal culture within-herd prevalence and milk ELISA herd sensitivity. Herd sensitivity increased as within-herd MAP prevalence increased; increasing from less than 20% in herds with 2% within-herd MAP prevalence to more than 95% in herds with 10% within-herd prevalence. Based on the low herd sensitivity in low-prevalence herds estimated from the current study, we inferred that whereas the milk ELISA herd test may be useful for a control program in a high-prevalence herd, as suggested by consensus recommendations (Collins et al., 2006), herd sensitivity is too low to detect a low-prevalence herd.

Herd sensitivity will increase as the ELISA cutoff is lowered, but this amplification will come at a cost to herd specificity (Martin et al., 1992). Although a selection of studies has used a 1-cow cutoff, the lower cutoff translates to a substantially lower herd specificity (Hendrick et al., 2005; Lombard et al., 2006). In the current data,

between 13 and 29% of the milk ELISA herd tests from MAP-negative herds had at least 1 positive ELISA result. For the purposes of applying the milk ELISA test in MAP control programs, this number of false-positive herd results was deemed to be too high. In the present study, a 2% positive cow cutoff was used because it allowed for an optimal balance between herd sensitivity and herd specificity, as well as for a greater extension to a variety of herd sizes.

Likelihood ratios (**LR**) were used to estimate the likelihood of a herd being MAP-positive, based on milk ELISA prevalence categories. In herds with a milk ELISA prevalence below 2%, the herd test result was non-informative. When the milk ELISA prevalence was > 4%, herds in this category were almost certainly MAP-positive. This information extends interpretation of herd milk ELISA data beyond a positive or negative herd result, and allows application of probability for a herd to be MAP-positive, based on the proportion of positive cow-level ELISA results.

The present results do not support the use of milk ELISA for detection of infected herds with < 2% within-herd MAP prevalence. A single herd sensitivity estimate for milk ELISA cannot be applied to all herds because herd sensitivity was significantly impacted by the within-herd MAP prevalence. Although herd specificity was high overall, the LR indicated that herds with greater than 0% but less than 2% milk ELISA prevalence were as likely to be a false-positive result as a true-positive result.

7.2.3 Herd-level Detection of MAP: Conclusions

In this study population, the sensitivity of both environmental culture and milk ELISA herd tests improved as apparent within-herd MAP prevalence increased. In herds with low within-herd MAP prevalence, the sensitivity of environmental culture was

substantially greater than the sensitivity of the milk ELISA, demonstrating that low-prevalence MAP-infected herds had a greater probability to be detected with environmental culture. In addition, reflecting the near perfect specificity of fecal culture (Collins et al., 2006), false-positive results from environmental culture are considered rare, and in this study population, only 1% of the environmental culture sets collected from the 18 MAP-negative herds were positive. In contrast, between 13 and 29% of the milk ELISA herd tests from MAP-negative herds had at least 1 positive ELISA result. Therefore, in this study population, environmental culture was preferred over milk ELISA for the detection of MAP-infected herds, as environmental culture better classified both low-prevalence, MAP-infected herds and MAP-negative herds.

7.3 Cow-level Detection of MAP

7.3.1 Cow-level Milk ELISA Agreement

In Chapter 4, agreement between 3 commercial milk ELISA kits was evaluated by way of agreement statistics, kappa values, and logistic models investigating factors that influenced positive milk ELISA agreement. Overall agreement was high (ELISA A/B: 0.985; ELISA A/C: 0.982; ELISA B/C: 0.992), and kappa values indicated moderate to substantial agreement (ELISA A/B: 0.60; ELISA A/C: 0.57; ELISA B/C: 0.79). Although agreement between negative test results of the 3 ELISAs was high (ELISA A/B: 0.992; ELISA A/C: 0.991; ELISA B/C: 0.996), agreement between positive test results was significantly lower (ELISA A/B: 0.61; ELISA A/C: 0.57; ELISA B/C: 0.79). Unless there is a specified testing strategy in place for the application of 2 ELISA tests to 1 milk sample, such as an initial screening (higher sensitivity) followed by a repeat diagnostic (higher specificity) test (Dohoo et al, 2009), the lower agreement between positive results

may result in discrepant results between 2 milk ELISA kits applied to the same milk sample.

For all 3 paired milk ELISA comparisons, fecal shedding of the MAP bacterium significantly increased the odds of test-positive milk ELISA agreement. The second predictor increasing the odds of test-positive milk ELISA agreement was the fecal culture within-herd MAP prevalence. This predictor was significant for test-positive agreement for 2 of the 3 pairs of ELISA kits compared. These results are consistent with consensus recommendations that milk ELISA is best applied in herds with known infection, high within-herd MAP prevalence, and clinical disease (Collins et al., 2006).

Although the overall and test-negative milk ELISA proportions of agreement were quite high, agreement on positive milk ELISA results was significantly lower. Odds of agreement between test-positive milk ELISA results increased with cow fecal shedding of MAP and higher within-herd MAP prevalence. Although a repeat ELISA may be more informative at a later time, as recommended by the Danish MAP control program (Nielsen, 2009), the low agreement on positive tests suggest there is limited gain using a different ELISA on the same milk sample to confirm a positive result.

7.3.2 Cow-level Milk ELISA Test Characteristics and Likelihood Ratios

Using 2 commercially available milk ELISA kits and cow-level fecal culture results, cow-level sensitivity and specificity of the milk ELISAs were estimated in Chapter 5, using both pseudogold standard methods and latent class analysis models. There was not a substantial difference between milk ELISA test characteristic estimates using pseudogold methods or latent class models. Sensitivity of ELISA A was 30.2% using a latent class model, and 28.4% using a pseudogold standard, while sensitivity of

ELISA B was 34.6% using a latent class model, and 33.1% using a pseudogold standard. Specificity of ELISA A was 99.2% using a latent class model, and 99.5% using a pseudogold standard, while specificity of ELISA B was 99.4% using a latent class model, and 99.7% using a pseudogold standard.

The pseudogold standard allows for comparison of results between the current study and other publications using this method of statistical evaluation. However, estimates using pseudogold standards have been questioned because they assume an accurate reference test, which is challenging to achieve in paratuberculosis research (Nielsen et al., 2002; Wang et al., 2011). The latent class model allows for estimation of test characteristics without necessitating a decision regarding disease status of the animal. Both methods were presented because each is relevant, and it was not an objective to select 1 statistical method as preferential over the other.

Accurate specificity estimates of the commercial ELISA kits are essential tools in a herd MAP control program, because producers and veterinarians must know the risk of false-positive results. This is particularly important when a positive test can result in culling of the animal. Specificity of paratuberculosis tests can be difficult to determine because the latent period of the disease prevents accurate use of a reference test to establish absence of infection in a cow. Therefore, the best method to define absence of infection in a cow is to establish a herd as non-infected, and then evaluate specificity from cows within non-infected herds (Collins et al., 2005). The repeated herd tests to establish herds as MAP-negative in this dataset allowed for accurate estimates of specificity.

One method to extract information from ELISAs, beyond a dichotomous ELISA result, is the development of LRs, which were assessed in Chapter 5. Categorical LRs for both ELISA A and ELISA B were significantly different between the lowest and uppermost categories of ELISA results. Practically, this means that for a cow with an ELISA result in the upper category, her likelihood of being MAP-infectious is very high, and a confirmatory test such as fecal culture would provide very little additional information regarding her MAP disease status. Furthermore, this cow should either be managed carefully (e.g. colostrum discarded, no contact with young stock), or culled for slaughter. The particular recommendation will depend on the specific objectives of the herd's MAP control program. Likely due to a small sample size, there was no significant difference between LRs for middle ELISA categories. If this study were repeated with a larger sample size, a more clear and distinct direct relationship between mid-categories and probability of infectiousness could be seen.

The establishment of test characteristics for 2 commercial milk ELISAs provides valuable information for practitioners and producers interpreting results from these kits. Test characteristics estimated using pseudogold standard methods and latent class analyses were not different. Likelihood ratios extend the use of ELISA results beyond a dichotomous diagnosis, and our results suggest that cows with ELISA results in the upper ranges of quantitative values are almost certainly MAP-infectious. Likelihood ratios allow more specific management strategies to be implemented for these cows.

7.3.4 Repeated Cow-level Milk ELISA Testing

The focus of Chapter 6 was to evaluate the additional diagnostic information gained in a repeat milk ELISA test. Sensitivity and specificity of initial tests, and initial

and repeated tests combined, were evaluated. For a 6 mo test interval, observed Se of the milk ELISA was 22.0% (95% CI: 15.6-28.3%) and 32.6% (95% CI: 26.3-41.2%), for initial and combined initial and repeated tests, respectively. Specificity of an initial ELISA was 99.6% (95% CI: 99.4-99.8%), and 99.2% (95% CI: 98.8-99.5%) for combined tests. For a 12 mo interval, Se of the milk ELISA was 25.6% (95% CI: 18.2-33.0%) and 45.3% (95% CI: 35.6-54.3%) for initial and combined initial and repeated tests, respectively. Specificity of an initial ELISA was 99.6% (95% CI: 99.4-99.9%), and 98.9% (95% CI: 98.4-99.4%) for combined tests.

Combined information from initial and repeated milk ELISA tests increased the probability of detecting a MAP-infected cow, compared to an initial ELISA test only. For 6 mo test intervals, and especially 12 mo test intervals, a repeated ELISA provided considerable gain in sensitivity with a relatively small decrease in specificity. The impact of a repeated test on specificity was minimal.

Logistic regression models were determined to identify significant predictors for a MAP-infected cow, initially milk ELISA-negative, to test positive on a repeat milk ELISA test. The probability to detect a MAP-infected cow using a 12 mo test interval tended ($P = 0.09$) to be higher compared to a 6 mo interval. An annual test, performed at a pre-determined point in a cow's lactation cycle, and using a milk sample that is already being collected as part of routine herd milk testing, would be acceptable to many producers from both a convenience and cost perspective. Lack of significance of the test interval at $P \leq 0.05$ may have been related to the relatively small number of MAP-infected cows in the dataset.

The quantitative value (S/P ratio (%)) of an initial negative milk ELISA test was predictive of a MAP-infected cow being detected on a repeat milk ELISA test. Results suggest it is beneficial to consider the magnitude of the ELISA value, in addition to the dichotomous result. If, for example, a producer is unable to repeat testing on all ELISA negative cows, the quantitative value may be useful to target repeat testing to cows with a value close to the cutoff.

In our dataset, there were 16 cows with a positive milk ELISA result on both initial and repeat tests. Small sample size precluded application of statistical models to this subgroup. However, it is likely that cows with repeated positive results are more likely to be infected with MAP and therefore should be managed as such in a herd control program.

Repeated testing in a MAP-positive herd provides an improved understanding of a cow's MAP status, and reduces risk of misclassification based on a single test result. A repeat milk ELISA test improves the probability to detect a MAP-infected cow, with minimal loss of specificity. Using a 12 mo test interval provides a greater increase in sensitivity than a 6 mo interval, relative to a single initial test. Cows with repeated positive ELISA results are more likely to be MAP-infected and should be a priority for risk management strategies.

7.3.5 Cow-level Detection of MAP: Conclusions

For the cow-level detection of MAP, test characteristics of commercial milk ELISA kits were similar to estimates previously published for in-house milk ELISA kits. Data analyses demonstrated that the quantitative value of a cow's milk ELISA result is useful information when making cow management choices within a herd MAP control

program. As well, repeated milk ELISA testing can substantially increase the probability of detecting a MAP-infected cow. Making use of milk ELISA results, beyond a dichotomous positive or negative diagnosis, will improve the quality of information gained from cow-level milk ELISA test results, and could increase the success of control programs within MAP-infected herds.

7.4 Future Research Directions

This evaluation of herd-level and cow-level diagnostic tests for detection of MAP provides new information to assist in the application of test strategies to herds motivated to control paratuberculosis. The research also allows for identification of unanswered questions that should be the focus of future research programs, in order that the dairy industry will continue to realize improved management and control of paratuberculosis.

It would be beneficial to increase the sensitivity of herd paratuberculosis tests, in particular for the detection of low-prevalence herds. If environmental manure samples are already being collected, a second sample that could easily be gathered from a herd is a bulk tank milk sample. Currently, work is being done to optimize PCR techniques for bulk tank milk detection of MAP. The bulk tank milk sample would be more convenient and efficient to collect than individual milk samples from the entire herd, and would be more economical to process in the laboratory. Conceivably, performance of PCR on bulk tank milk samples is also impacted by low within-herd prevalence. A lower concentration of MAP in the bulk tank is expected in a low-prevalence herd, making detection challenging. It would be ideal to evaluate the efficacy of combining environmental culture and bulk tank milk PCR, in order to assess if this testing combination could increase sensitivity to detect low-prevalence herds. Bulk tank milk ELISA alone has been

examined (Nielsen et al., 2000; van Weering et al., 2007). While bulk tank milk ELISA has limited sensitivity as a stand-alone test, it might have merit as part of a suite of tests to maximize sensitivity.

Results of our analyses indicated a trend for manure storage areas and cow manure concentration areas to be more often culture-positive than samples from sick cow or maternity pens. As well, many farms did not meet the criteria for collection in sick cow or maternity pens (i.e. having 2 or more cows in the pen and no manure clean-out between animals). It should be determined if a sampling protocol focused on manure storage and cow manure concentration areas could be used, without a loss in test sensitivity. In addition, an evaluation of direct PCR for detection of MAP in environmental samples would be beneficial. Because PCR is less costly and more time effective than fecal culture (Tiwari et al., 2006), it would benefit the dairy industry to determine if PCR as a herd diagnostic test was as sensitive and specific as culture of environmental samples.

The sensitivity of environmental culture in tie-stall facilities remains unknown. In the current study population, only 1 MAP-positive herd had tie-stall facilities, precluding statistical evaluation of the relationship between housing type and environmental culture sensitivity. A substantial proportion of the dairy industry, particularly in Quebec (CDIC, 2012), utilizes this type of housing. In a tie-stall barn, cattle movement and therefore, mixture of manure, is minimal. Without this cow movement, in a tie-stall, it may be more challenging to collect composite manure samples representing many cows, as compared to a free-stall facility. Therefore, it is important to evaluate the probability that environmental culture can detect MAP-positive herds housed in tie-stall facilities.

The application of repeated milk ELISA tests to improve the probability to detect MAP-infected cows appears promising. It would be beneficial to determine the fluctuations in antibody response throughout the lactation cycle of a MAP-infected cow. This would allow for application of a repeated testing program at a point in lactation when the ELISA response is greatest. Finally, an economic analysis to investigate the economic benefits and costs of repeated testing is necessary, as well as a determination of the testing interval that maximizes return on investment. This will enable veterinarians to better guide producers to develop successful and sustainable MAP control programs.

7.5 References

Canadian Dairy Information Centre. 2012. Dairy barns by type in Canada 2012. Accessed November 18, 2013. http://www.dairyinfo.gc.ca/index_e.php?s1=dff-fcil&s2=farm-ferme&s3=db-el.

Chiodini, R. J., H. J. Van Kruiningen, and R. S. Merkal. 1984. Ruminant paratuberculosis (Johne's disease): The current status and future prospects. Cornell Vet. 74:218-262.

Christensen, J. and I. A. Gardner. 2000. Herd-level interpretation of test results for epidemiologic studies of animal diseases. Prev. Vet. Med. 45:83-106.

Collins, M. T., I. A. Gardner, F. B. Garry, A. J. Roussel, and S. J. Wells. 2006. Consensus recommendations on diagnostic testing for the detection of paratuberculosis in cattle in the United States. J. Am. Vet. Med. Assoc. 229:1912-1919.

Collins, M. T., S. J. Wells, K. R. Petrini, J. E. Collins, R. D. Schultz, and R. H. Whitlock. 2005. Evaluation of five antibody detection tests for diagnosis of bovine paratuberculosis. Clin. Diagn. Lab. Immunol. 12:685-692.

Dohoo, I. R., S. W. Martin, and H. Stryhn. 2009. Veterinary Epidemiologic Research. 2nd ed. VER, Inc., Charlottetown, P.E.I., Canada.

Dorshorst, N. C. and J. E. Lombard. 2006. Decision analysis model for paratuberculosis control in commercial dairy herds. Prev. Vet. Med. 75:92-122.

Garry, F. 2011. Control of paratuberculosis in dairy herds. Vet. Clin. North Am. Food Anim. Pract. 27:599-607.

Hendrick, S., T. Duffield, K. Leslie, K. Lissemore, M. Archambault, and D. Kelton. 2005. The prevalence of milk and serum antibodies to Mycobacterium avium subspecies paratuberculosis in dairy herds in Ontario. Can. Vet. J. 46:1126-1129.

Kalis, C. H., M. T. Collins, H. W. Barkema, and J. W. Hesselink. 2004. Certification of herds as free of Mycobacterium paratuberculosis infection: Actual pooled faecal results versus certification model predictions. Prev. Vet. Med. 65:189-204.

Lombard, J. E. 2011. Epidemiology and economics of paratuberculosis. Vet. Clin. North Am. Food Anim. Pract. 27:525-35.

Lombard, J. E., T. M. Byrem, B. A. Wagner, and B. J. McCluskey. 2006. Comparison of milk and serum enzyme-linked immunosorbent assays for diagnosis of Mycobacterium avium subspecies paratuberculosis infection in dairy cattle. J. Vet. Diagn. Invest. 18:448-458.

Martin, S. W., M. Shoukri, and M. A. Thorburn. 1992. Evaluating the health status of herds based on tests applied to individuals. Prev. Vet. Med. 14:33-43.

Nielsen, S. S. 2009. Use of diagnostics for risk-based control of paratuberculosis in dairy herds. In Pract. 31:150-154.

Nielsen, S. S., C. Gronbaek, J. F. Agger, and H. Houe. 2002. Maximum-likelihood estimation of sensitivity and specificity of ELISAs and faecal culture for diagnosis of paratuberculosis. Prev. Vet. Med. 53:191-204.

Nielsen, S. S., S. M. Thamsborg, H. Houe, and V. Bitsch. 2000. Bulk-tank milk ELISA antibodies for estimating the prevalence of paratuberculosis in Danish dairy herds. Prev. Vet. Med. 44:1-7.

Pillars, R. B., D. L. Grooms, and J. B. Kaneene. 2009. Longitudinal study of the distribution of Mycobacterium avium subsp. paratuberculosis in the environment of dairy herds in the Michigan Johne's disease control demonstration herd project. Can. Vet. J. 50:1039-1046.

Raizman, E. A., S. J. Wells, S. M. Godden, R. F. Bey, M. J. Oakes, D. C. Bentley, and K. E. Olsen. 2004. The distribution of Mycobacterium avium ssp. paratuberculosis in the environment surrounding Minnesota dairy farms. J. Dairy Sci. 87:2959-2966.

Smith, R. L., Y. H. Schukken, A. K. Pradhan, J. M. Smith, R. H. Whitlock, Van Kessel J.S., D. R. Wolfgang, and Y. T. Grohn. 2011. Environmental contamination with Mycobacterium avium subsp. paratuberculosis in endemically infected dairy herds. Prev. Vet. Med. 102:1-9.

Tiwari, A., J. A. VanLeeuwen, S. L. McKenna, G. P. Keefe, and H. W. Barkema. 2006. Johne's disease in Canada part I: Clinical symptoms, pathophysiology, diagnosis, and prevalence in dairy herds. Can. Vet. J. 47:874-882.

USDA. 2010. Uniform Program Standards for the Voluntary Bovine Johne's Disease Control Program. USDA-APHIS-VS, National Animal Health Monitoring System, Fort Collins, CO (APHIS 91-45-016).

van Weering, H., G. van Schaik, A. van der Meulen, M. Waal, P. Franken, and K. van Maanen. 2007. Diagnostic performance of the Pourquier ELISA for detection of antibodies against Mycobacterium avium subspecies paratuberculosis in individual milk and bulk milk samples of dairy herds. Vet. Microbiol. 125:49-58.

Wang, C., B. W. Turnbull, S. S. Nielsen, and Gröhn Y.T. 2011. Bayesian analysis of longitudinal Johne's disease diagnostic data without a gold standard test. J. Dairy Sci. 94:2320-2328.

Wells, S. J., S. M. Godden, C. J. Lindeman, and J. E. Collins. 2003. Evaluation of bacteriologic culture of individual and pooled fecal samples for detection of Mycobacterium paratuberculosis in dairy cattle herds. J. Am. Vet. Med. Assoc. 223:1022-1025.

Whitlock, R. H. 2010. Paratuberculosis control measures in the USA. Pages 319-329 in Paratuberculosis: organism, disease, control. M. A. Behr and D. M. Collins ed. CAB International, Wallingford, UK.